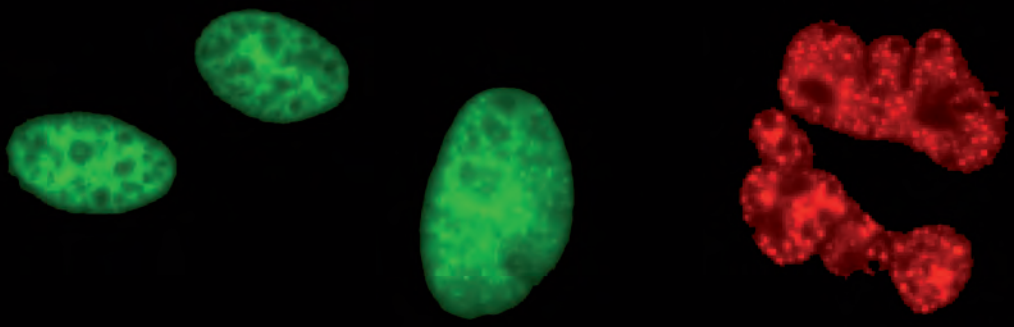
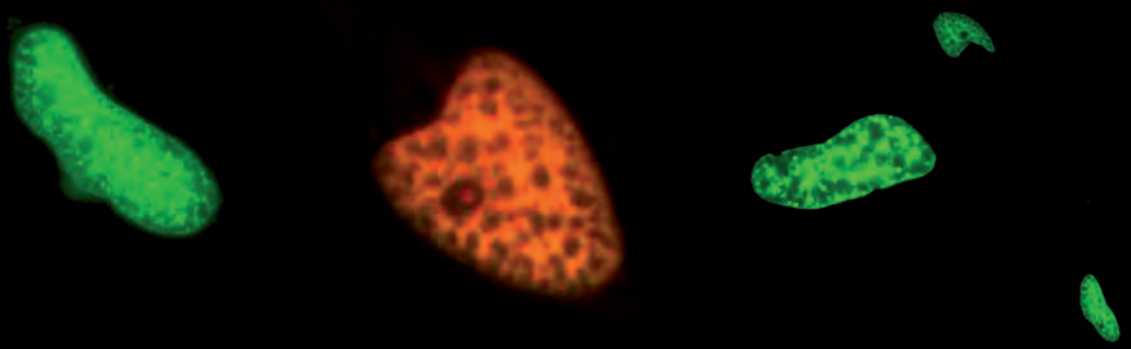


Clinical and Biological Insights into the Novel Epigenetic Regulator TET2



Mariam Aslanyan

Stellingen

1. At the start of this PhD project, there were 0 papers published on the gene/protein TET2. Shortly before this thesis was sent out for printing, a Pubmed search for TET2 generated 532 hits!

2. TET2 is trending...

3. Associating genetic mutations with patient survival is crucial, but it is merely the beginning of a very long road towards more sophisticated, targeted treatments.

4. It is my observation that the superficiality that seems to be taking over the world has reached the scientific community, and much more attention and credit is being given nowadays to finding new disease mutations, rather than taking that long road.

5. "I realized recently that if I examine it in a day-to-day sense, I have one job in science. It's not curing malaria, which is what my grant says it should be. My job, in essence, is to move small amounts of liquid from one place to another. That's it." (Adam Ruben, Science website, Experimental error series)

6. "The most exciting phrase to hear in science, the one that heralds new discoveries, is not, "Eureka!" ("I found it!") but rather, "Hmm... that's funny..." (Isaac Asimov)

7. "We are, quite literally, star dust." (Christof Koch, Consciousness: Confessions of a Romantic Reductionist)

8. "Ultimately, though, pinning a variant to a function is the part of science least amenable to high-throughput approaches. You can easily generate data that now demands years' worth of work to digest." (Mark de Pristo)

9. Research is driven by pre-postulated hypotheses. Often, in our pursuit to prove these hypotheses, we forget that disproving them is just as important.

10. It is rather arrogant to assume that most of our hypotheses on the works of Nature will prove to be correct, when the matter of fact remains that modern day scientists are attempting to decipher over 4 billion years of evolution.

11. "One's ideas must be as broad as Nature, if they are to interpret Nature" (Sherlock Holmes, A Study In Scarlet, Arthur Conan Doyle)

12. One CAN spend the happiest hours of their PhD in the dark….room, where the microscope is located!

Clinical and Biological Insights into the Novel Epigenetic Regulator TET2

Mariam Aslanyan

Clinical and Biological Insights into the Novel Epigenetic regulator TET2

Mariam Aslanyan

Cover Design: Mariam Aslanyan

Thesis layout and design: Off Page, Amsterdam

Print: Off Page, Amsterdam

ISBN: 978-94-6182-549-0

The studies described in this thesis were performed from September 2008 till December 2012 at the Laboratory of Hematology, Department of Laboratory Medicine of the Radboud University Medical Centre Nijmegen (presently: Radboudumc), Nijmegen, The Netherlands.

Clinical and Biological Insights into the Novel Epigenetic Regulator TET2

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 30 april 2015
om 16.30 uur precies

door

Mariam Grigor Aslanyan

geboren op 6 januari 1983
te Sofia, Bulgarije

Promotor:
Prof. dr. J.H. Jansen

Copromotor:
Dr. B.A. van der Reijden

Manuscriptcommissie:
Prof. dr. A.H.M. Geurts van Kessel
Dr. G.A. Huls
Prof. dr. J.J. Schuringa (UMCG)

На моето мило семейство

TABLE OF CONTENTS

Chapter 1

Introduction

Parts of this section have been previously published:

* *Langemeijer SMC, Kuiper RP, Berends M, et al.: Acquired mutations in TET2 are common in myelodysplastic syndromes. Nature Genetics 41:838-U102, 2009*

** *Langemeijer SM, Aslanyan MG, Jansen JH: TET proteins in malignant hematopoiesis. Cell Cycle 8:4044-8, 2009*

*** *S.M.C. Langemeijer, M.G. Aslanyan en J.H. Jansen: TET2 mutaties komen frequent voor bij myeloïde maligniteiten. Ned Tijdschr Hematol 2010;7:131, 2010*

1. TET2 MUTATIONS IN MYELOID NEOPLASMS

1.1 Discovery of TET2 mutations in MDS and MPN

In the past decade, novel technological developments, such as SNP arrays and next-generation sequencing, have allowed researchers to gain better insight into the genetic background of myeloid malignancies. Consequently, a number of previously undetectable recurrent somatic alterations have been identified and their repercussion on clinical outcome of patients has been extensively studied resulting in improved risk stratification.

In 2009, two original reports were published implicating for the first time the *TET2* gene in the pathogenesis of myelodysplastic syndromes (MDS) and myeloproliferative neoplasms (MPN), respectively [1,2]. Due to the limitations of traditional karyotyping and FISH analysis, gross chromosomal aberrations (such as -5q, -7q, -7,) could only be detected in approximately half of the MDS patients, leaving 50% of the cases unexplained (box 1). In addition, mutations in *NRAS* (15%), *TP53* (10%), *RUNX1* (10%), and *FLT3* (5%) had been described. Using 250k SNP array-based genomic profiling, we identified a microdeletion encompassing a minimal region on chromosome 4q24 containing only the *TET2* and *PPA2* genes. Sanger sequencing revealed that the *PPA2* gene was unaffected, but the remaining copy of *TET2* carried an acquired nonsense mutation introducing a premature stop codon in the midst of the protein. Subsequent analysis of the rest of the MDS patient cohort showed that nonsense, missense and frameshift *TET2* mutations, scattered along the whole coding region of the gene, were present in 26% of the MDS patients with/without 4q24 abnormalities. In several subjects, more than one *TET2* mutation was detected suggesting biallelic inactivation. Afflicting a quarter of the patients, *TET2* became the most frequently mutated gene in MDS thus far [1].

Based on a report in 2005 [3] describing six MDS/AML patients with acquired rearrangements on 4q24, Delhommeau et al. [2] further pursued the potential implication of this chromosomal region in leukemogenesis. Employing an alternative approach, the authors discovered involvement of the 4q24 region in patients with myeloprolifertive neoplasms (MPN), as well as MDS and AML. A large percentage of MPN patients (box 2) harbor a common genetic defect in the Janus kinase (JAK) 2 gene, namely the *JAK2* V617F activating mutation. Evidence from mouse models and patients had suggested that the *JAK2* defect alone cannot always account for the progression of the disease. This led the authors to the hypothesis that an additional early genetic event may arise in the hematopoietic stem cells which could be common to MPN, MDS and AML. Similar to our study, they also reported truncating single or double mutations affecting the whole coding portion of the *TET2* gene in an overall 15% of patients with diverse myeloid malignancies. Fractionation of the CD34+ population of progenitor cells into CD34+CD38- (immature HSCs) and CD34+CD38- (more committed cells) indicated that *TET2* mutations were present in

a small number of stem cells and increased in the committed progenitor population, implying that *TET2* defective cells have a proliferative advantage. To prove that further, CD34+ MPN patient cells carrying a *JAK2* V617F mutation either with a *TET2* or without a *TET2* aberration were engrafted into NOD-SCID mice. *TET2* positive cells had an increased proliferative capacity and eventually a clone carrying only a *TET2*, but no *JAK2* mutation, outgrew and completely dominated the bone marrow skewing blood cell reconstitution towards the myeloid lineage.

Box 1

MDS represent a heterogeneous group of clonal hematopoietic stem cell disorders characterized by dysplasia of one or more of the myeloid blood cell lineages (erythroid, granulocytic, monocytic and/or megakaryocytic) (Fig.1). It is among the most prevalent hemopathies affecting the elderly population with 60-65 years being the median age of onset, occurring more often in males than in females. Patients suffer from peripheral blood cytopenias and have an increased risk of progression to acute myeloid leukemia (AML). Although about 30% of the cases progress to AML, most MDS patients eventually die of complications arising from bone marrow failure. MDS can be very difficult to diagnose since the initial symptoms are not specific to the disease. At present, diagnosis is based on careful cytomorphologic examination of the bone marrow, cytogenetic analysis, molecular studies and recently, flow cytometry has become a powerful tool in discriminating MDS from other disorders. Depending on the type of MDS and the age of the patient, treatment options include palliative care, erythroid stimulating agents (ESAs), lenalidomide and hypomethylating agents. Hematopoietic stem cell transplantation (HSCT) remains the sole curative option but its use is restricted to younger patients due to the age-associated morbidity and mortality.

Box 2

MPN are clonal bone marrow stem cell disorders characterized by an expansion of mature cells from the myeloid, erythroid or megakaryocytic lineages. Classic MPNs comprise: polycythemia vera (PV), essential thrombocythemia (ET) and primary myelofibrosis (PM). Although diagnosed as distinct entities PV, ET and PM show considerable overlap in terms of morphology, clinical presentation and symptoms. Interestingly, all three types of MPNs also share a common genetic defect within the JAK2 kinase, the activating JAK2 V617F mutation (~95% of PV, 50-70% ET, and 40-50% PM)[4] which renders the protein in a constitutively active state. As for MDS, the treatment for MPN is mostly supportive. At present, the most promising new therapeutics available for MPN patients are JAK2 inhibitors.

*1.2 TET2 mutations in other myeloid malignancies_*

Following the discovery of *TET2* mutations in MDS and MPN, a number of groups addressed the question whether recurrent *TET2* defects were restricted to MDS/MPN, or occurred in a wide spectrum of myeloid disorders. This effort resulted in the addition of systemic mastocytosis (SM), chronic myelomonocytic leukemia (CMML), de novo AML, as well as secondary AML (sAML) to the list. *TET2* aberrations affected 29% of SM patients and segregated with *KIT* D816V mutation [5]. The highest frequency of *TET2* mutations however was reported in patients with CMML, where the incidence reaches up to 50% of the cases suggesting a particularly important role for TET2 in the pathogenesis of this disease. According to initial reports, the incidence of *TET2* mutations reported in AML ranged between 12-19%, and was higher in the case of sAML [7].

In contrast to adults, where TET2 represents a common pathogenic denominator, pediatric disorders seem to be driven by a distinct set of molecular events. *TET2* mutations are in fact rare in younger patients. *TET2* aberrations were absent in juvenile myelomonocytic leukemia (JMML) [8] and precursor B-cell ALL [9], and were only found in 3,8% of childhood AML cases [9]. Only one patient with childhood MDS/MPN has been described to harbor a *TET2* mutation so far [10].

Strikingly however, *TET2* mutations were recently found in normal elderly individuals with clonal hematopoiesis without signs of hematological malignancy [11]. This finding may indicate the presence of *TET2* mutations already at a pre-disease stage which could further progress towards distinct myeloid disorders depending on the additionally acquired genetic defects.

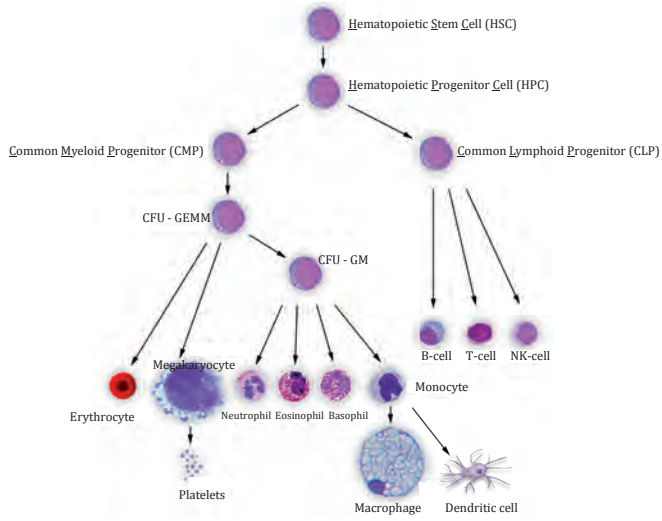
2. THE TET2 PROTEIN AND ITS (DYS) FUNCTION

2.1 Function of the TET proteins

The TET2 protein belongs to a family comprising TET1, TET2 and TET3. Long before TET2 entered the stage, TET1 had already made a short appearance. In 2003 Lorsbach et al. [12] reported TET1, i.e. the (10;11)(q22;q23) ten-eleven-translocation, thus giving the TET (ten-eleven-translocation) proteins their name, as novel fusion partner of the promiscuous MLL gene in a case of AML.

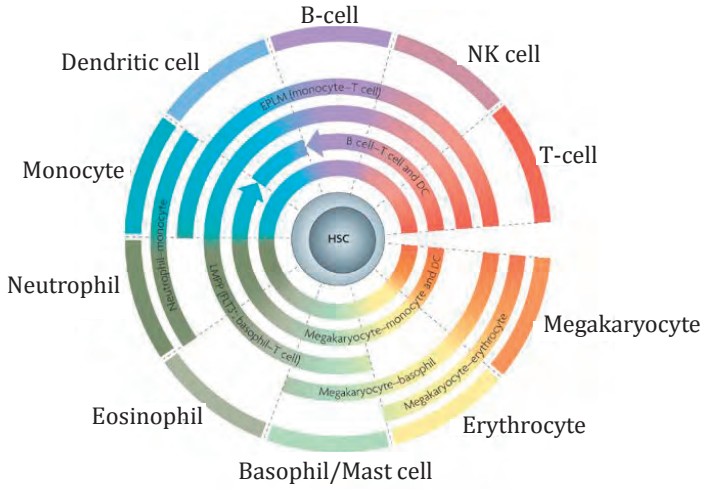
Employing a bioinformatics approach, Iyer et al. 13 predicted that the TETs belong to the superfamily of 2OGFeDO dioxygenases. All three members share a homologous C-terminal region, the 2OGFeDO (2-oxoglutarate and Fe(II)-dependent dioxygenase) catalytic domain responsible for their enzymatic activity. Unlike TET2, TET1 and TET3 have a CXXC domain in their N-termini. The DNA-binding CXXC domain present in the TET proteins occurs in several other chromatin modifiers, a prominent representative being the maintenance DNA-methyltransferase 1 (DNMT1) protein. Interestingly, TET2 lies in close proximity to

A



Adapted from W. Lensch, daley.med.harvard.edu

B



Adapted from Ceredig et al. [4]

Figure 1. *Hematopoiesis. Hematopoiesis (from Ancient Greek: haima blood; poiesis to make) is the process of blood cell formation. All mature blood cell types are progeny of the pluripotent hematopoietic stem cell (HSC). The conventional model of hematopoiesis represented in (A) has been widely accepted for over 30 years. It depicts blood cell development as a branching tree with two main trunks giving rise to either the CMP or the CLP. Myeloid cell production begins with the CFU-GEMM (colony forming unit that generates myeloid cells: granulocytes, erythrocytes, monocytes, megakaryocytes). The more restricted CFU-GM gives rise to granulocytes (neutrophils, eosinophils, basophils) and monocytes (which in turn can either become macrophages or myeloid-derived dendritic cells). B-cells,* ▶

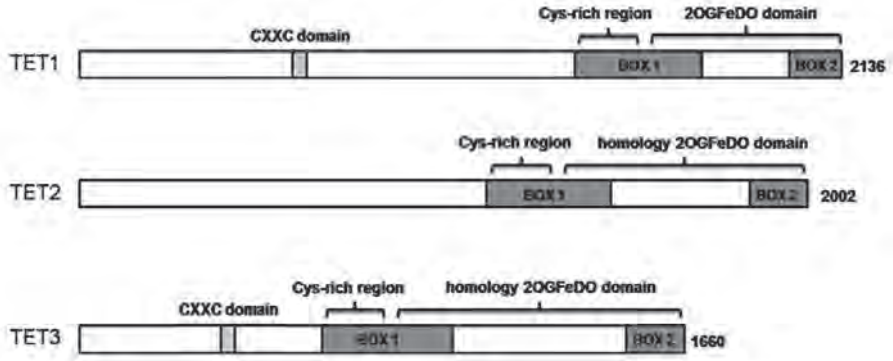


Figure 2. *Schematic representation of the TET proteins. The TET protein family consists of three members (TET1, TET2 and TET3) which share a homologous C-terminal region. The C-terminal portion encodes the catalytic domain of the TET proteins, the 2OGFeDO (2-oxoglutarate Fe(II)-dependent dioxynegase domain). Box 1 and box 2 regions are highly conserved between all three TET paralogs, as well as in TET orthologs. Unlike TET1 and TET3, TET2 does not contain a CXXC domain.*

the CXXC4 gene, implying that during evolution TET2 has lost its CXXC domain due to a local chromosomal rearrangement (Fig. 2).

Iterative sequence analysis showed that TET proteins are the mammalian homologues of the Trypanosome JBP1/2 proteins. JBP1 and JBP2 are DNA-modifying enzymes which catalyze the first step in the formation of base J (β-D-glucosylhydroxymethyluracil), a modified thymine, present in trypanosomes. Similar to 5-methylcytosine (5mC) in mammalian DNA, base J has been associated with gene silencing. A landmark paper in 2009 by Tahiliani et al. [14] showed that TET1 is able to modify 5-methylcytosine in mammalian DNA by converting it into 5-hydroxymethylcytosine (5hmc). Cells transfected with wild type TET1, but not a catalytic mutant form, displayed a decline in 5mC levels with a concomitant generation of 5hmC. In addition, *in-vivo* evidence demonstrated that the novel cytosine species is present in undifferentiated mouse embryonic stem cells (mESCs), and its levels decrease upon RNAi mediated depletion of endogenous TET1.

T-cells and NK (natural killer) cells are all derived from the CLP. However, recent models have challenged this apparent dichotomy and proposed a less strict separation between the various routes the HSCs can take. The cyclical hematopoiesis model (B) proposed by the group of Geoffrey Brown [4] suggests that intermediate progenitors have both myeloid and lymphoid potential. Abbreviations used are: CLP (common lymphoid progenitor); CMP (common myeloid progenitor); ELP (early lymphoid progenitor); GMP (granulocyte–monocyte progenitor); HSC (hematopoietic stem cell); LMPP (lymphoid-primed multipotent progenitor); MBP (myeloid–B-cell progenitor); MEP (myeloid–erythroid progenitor); MPP (multipotent progenitor); MTP (myeloid–T-cell progenitor).

Back-to-back with the study of Tahiliani et.al, another report appeared, showing that 5hmC is also present in very high levels in mouse cerebellar Purkinje neurons [15].

These findings led to the "re-discovery" of 5hmC as a normal physiological constituent of mammalian DNA after its original discovery in 1953 [16] remained largely neglected or failed to be reproduced by others, likely due to technical and methodological variations.

2.2 TET2 in the active DNA demethylation pathway

Fine tuning of gene expression is mediated via the intricate interplay between genetic and epigenetic regulatory mechanisms.

Note 1

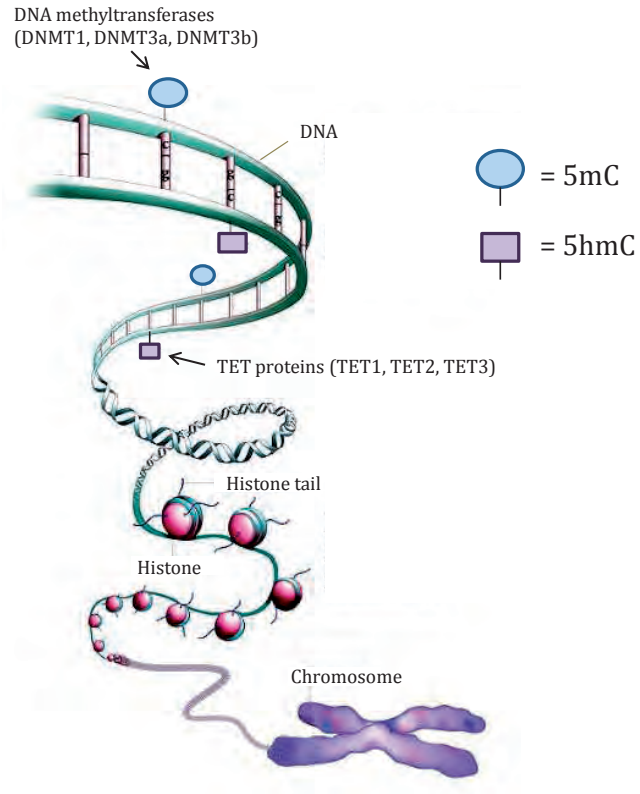
Epigenetics is the study of heritable changes in gene expression caused by mechanisms other than the underlying DNA sequence.

The epigenetic code consists of two main components: DNA methylation and histone modifications (Fig. 3). Whereas DNA methylation directly operates at the DNA level through the addition of a methyl group (-CH_3) at the 5th position of the cytosine pyrimidine nucleotide, histone modifications occur at the "tails" of histone molecules around which DNA is tightly wrapped. Together they form nuclear chromatin. Open chromatin allows for easy access of the RNA Pol II transcription machinery, hence signifying active gene expression, while compact chromatin ensures gene silencing.

DNA methylation is a crucial epigenetic mark associated with transcriptional repression, which has a broad range of regulatory roles in diverse biological processes. Cytosine methylation occurs prominently at promoter regions of genes within CpG islands (box 3).

Box 3

CpG dincleotides form CpG islands (CGIs). CGIs are short interspersed DNA sequences with an average length of 1000 base pairs which have an elevated G+C base composition (at least 55% GC content). Most, perhaps even all, are sites of transcription initiation and approximately 70% of the currently annotated gene promoters are associated with CGIs making them the most common promoter type in vertebrate genomes. CGIs are important regulatory structures where local changes of DNA and chromatin composition strongly influence transcription activation and repression. The CGIs of most genes are devoid of DNA methylation. CGIs may become methylated during normal development which leads to permanent silencing of the associated promoter [18-20]. Abnormal promoter methylation is among the hallmarks of cancer.



Adapted from Qiu [17]

Figure 3. *Components of epigenetic regulation. Cytosine methylation is mediated via the DNA-methyltransferases (DNMT1, the maintenance DNMT, and DNMT3a and DNMT3b, the de novo DNMTs). DNA methylation represses gene expression. The TET proteins (TET1, TET2, and TET3) further convert 5mC into 5-hydroxymethylcytosine (5hmC), an intermediate in the active DNA demethylation pathway. DNA is wrapped around histones to form nucleosomes. Each nuclesome consists of ~147 DNA bases wound around a histone octamer containing 2 copies of each of the core histones, H2A, H2B, H3 and H4, as well as ~80bp of linker DNA and the H1 linker histone. Histones have N-terminal "tails" which are subject to a variety of modifications, such as acetylation, methylation, phosphorylation, ubiquitination and sumoylation. These modifications result in complex structural and functional alterations which either repress or activate the associated genes. DNA together with the histone proteins forms a higher order structure called chromatin, whose further compaction results in the formation of chromosomes.*

In most mammals, 5mC is required for allele-specific expression of imprinted genes, transcriptional repression of retrotransposons, and X-chromosome inactivation in females [21]. Disturbances in 5mC signatures are among the hallmarks of cancer initiation and progression. DNA methylation and the proteins which

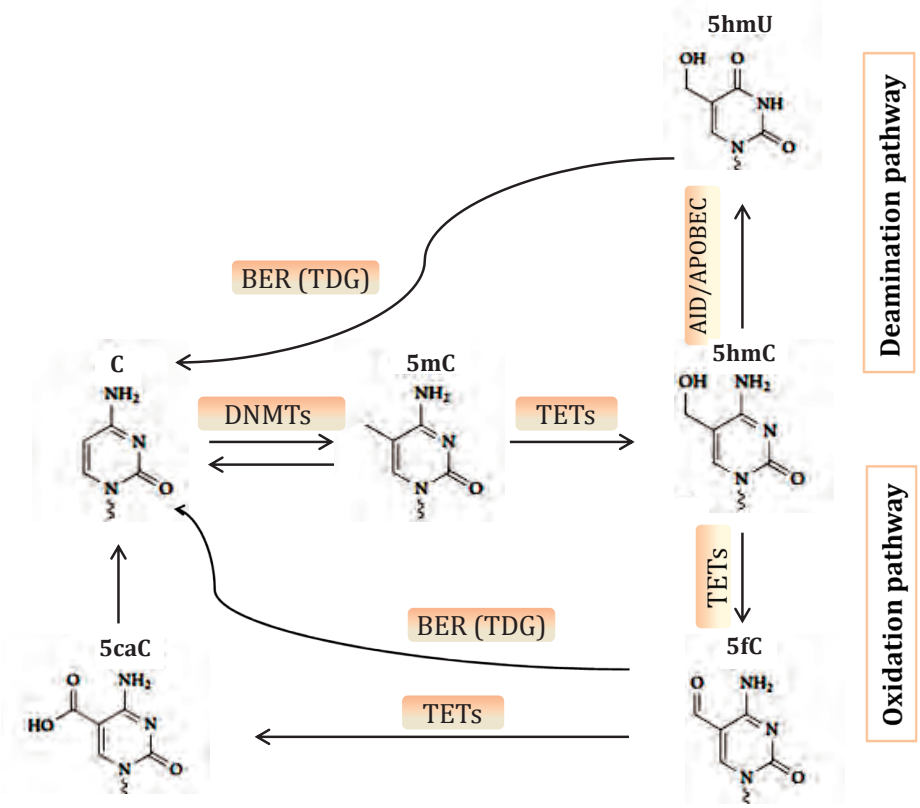
orchestrate this process have been the subject of extensive research. Once *de novo* methylation has been established by the DNA-methyltransferases 3a (DNMT3a) and 3b (DNMT3b), the pattern is faithfully maintained through cell divisions by DNMT1. DNMT1 is recruited to replication foci during S-phase of the cell cycle, where it preferentially binds hemi-methylated DNA to copy methylation marks onto the daughter DNA strands [22].

DNA methylation must be kept in balance for cells to function normally from the very conception of life. Mammalian germ cells undergo dynamic global re-patterning of methylation signatures at two stages: upon fertilization in the zygote, and in primordial germ cells (PGCs), which are direct descendents of the sperm and oocyte. Currently, it is considered that these dramatic waves of demethylation and re-methylation allow the zygote to erase the epigenetic signatures inherited from the gametes (with the exception of parental imprinting) and regain a state of developmental totipotency. Notably, the loss of methylation occurring in the paternal pronucleus has been firmly linked to TET3 in the last few years [23,24]. Parental imprint erasure, and subsequent re-establishment of genomic imprinting as to correspond the gender of the embryo in PGCs, also plays a central role in the restoration of developmental potential [25].

Two main mechanisms exist for reversion of DNA methylation. Passive DNA demethylation is achieved by simply blocking methylation of the newly synthesized DNA strand during replication which results in the replication-dependent dilution of DNA methylation. Although several potential mechanisms for active DNA demethylation had been proposed, the process remained largely a mystery, as did the proteins responsible for it. An active DNA demethylation pathway could involve oxidative deamination of 5mC, which would result in the generation of a thymine, creating a guanine: thymine (G: T) mismatch on DNA. This would trigger the activity of the base excision repair (BER) machinery thus restoring unmethylated C, at the place of 5mC. The search for the mammalian DNA demethylase yielded a number of potential candidates, including the thymine DNA glycosylase (TDG) and the Growth Arrest and DNA Damage (GADD45a) and GADD45b among others. However, evidence for the involvement of these proteins in active DNA demethylation remained controversial.

The discovery of the TET proteins and the product of their enzymatic activity, 5hmC, put active DNA demethylation back in the limelight. Essentially two main mechanisms have been proposed involving the TETs and 5hmC: the deamination pathway and the oxidation pathway (Fig. 4).

Mice lacking the activation-induced cytosine deaminase (AID) exhibit significant defects in genome-wide DNA demethylation in primordial germ cells [26]. A report from Bhutani et al. [27], demonstrated that promoter demethylation of the pluripotency genes OCT4 and NANOG during in vitro nuclear reprogramming was facilitated by AID. Notwithstanding, AID is unable to operate on dsDNA and deamination of 5mC



Adapted from Branco et al. [36]

Figure 4. *Active DNA demethylation pathways. Active DNA demethylation has been suggested to occur via two alternative mechanisms: the deamination and the oxidation pathway. 5hmC can be deaminated by the AID/APOBEC enzymes to generate 5-hydroxymethyluracil, 5hmU. Alternatively, 5hmC is oxidized by the TET proteins to its formyl-, and carboxyl- derivatives, 5fC and 5caC. 5hmU, 5fC and 5caC are all recognized as mismatched bases by the base excision repair (BER) machinery and excised by the thymine DNA glycosylase (TDG).*

in vitro occurs at a very low rate as compared to unmethylated C [28,29].These results have raised the question, whether the involvement AID in active DNA demethylation occurs directly, or it requires an intermediate. In 2011 Guo et.al [30] provided evidence that upon overexpression of TET1 in HEK293 cells, there was a prominent increase of 5-hydroxymethyluracil (5hmU), the product of 5hmC deamination. They suggested a model in which the TET proteins would first generate 5hmC through oxidation of 5mC, which would subsequently be deaminated by the AID/APOBEC deaminases to yield 5hmU. Eventually, 5hmU would be excised by the BER machinery and an unmethylated C would be restored in DNA.

Shortly after these findings, an alternative pathway of DNA demethylation was reported by two independent groups. Ito et al. [31] hypothesized that 5mC could be oxidized first to 5hmC, and then further to the aldehyde 5-formylcytosine (5fC), and carboxylic acid, 5-carboxylcytosine (5caC), by the TET proteins. Using thin-layer chromatography (TLC) and mass spectrometric (MS) methods, they were indeed able to show the presence of all three oxidized cytosine species, 5hmC, 5fC and 5caC, in genomic DNA of mouse ESCs. Knockdown of Tet1 in mESCs led to a decrease of all three nucleotides. Based on the same surmise, He et al. [32], demonstrated that purified TET2 incubated with 5mC and 5hmC substrates generated a product which eluted in HPLC at the same time as a chemically synthesized 5caC standard. DNA of HEK293T cells transfected with wild-type TET2, but not a catalytic mutant, also contained the oxidation derivative. Furthermore, incubation of TDG with 5hmC and 5caC –containing DNA substrates, showed that TDG was able to excise 5caC but not 5hmC. Ectopic expression of TDG in HEK293T cells diminished the amount of 5caC generated by the co-transfected TET2 protein, while not affecting 5hmC levels. Depletion of TDG in mESCs resulted in reduced 5caC excision activity. Maiti and Drohat [33] provided biochemical evidence that TDG recognizes and rapidly excises 5fC and 5caC, yet it cannot remove 5hmC.

Interestingly, a report by Nabel et al. [34] has challenged the plausibility of the deamination pathway. Purified AID/APOBEC had low deamination activity on 5mC as compared to unmodified cytosine, and no detectable deamination activity towards 5hmC. Furthermore, upon overexpression of AID/APOBEC, 5hmU was not found in genomic DNA, whereas the oxidation intermediates remained detectable. Further support of this notion came from another study [35] showing that AID was not capable of deaminating 5hmC likely due to the size of the modified cytosine i.e. AID's enzymatic activity is inversely proportional to the electron cloud size of C5-cytosine making it an unlikely candidate in 5hmC removal from DNA. Hence, biochemical and cell based assays have confronted the feasibility of the deamination step.

2.3 Genome-wide distribution of TET/ 5hmC and potential functions of DNA hydroxymethylation

The discovery of 5hmC posed somewhat of a technical puzzle, when it became clear that classical methods for detection of 5mC, such as bisulfite treatment and methylation sensitive restriction digestions [37,38] do not distinguish between 5mC and 5hmC. New tools were required in order to address the position and function of 5hmC in mammalian DNA.

A rapid succession of studies has accumulated mapping the position of 5hmC, and recently 5fC, in the genome and their putative role in the regulation of gene expression. Since these modifications represent < 1% of the nucleotides, as

compared to the normal DNA constituents (A, T, C and G), highly selective and sensitive methods for detection and quantification have been developed within a short time. Initially, chromatin immunoprecipitation (ChIP)-sequencing based approaches were employed, either making use of antibodies recognizing 5hmC itself, or by pulling down of 5hmC after chemical conversion.

The groups of Helin and Zhang performed mapping of both TET1 and 5hmC in mouse ESCs [39,40]. TET1 was detected with highest density around the transcription start sites (TSS) and within the gene bodies of 6,573 genes [39]. Both studies confirmed the presence of TET1 at the TSS of pluripotency related genes and a highly significant overlap with bivalent domains regulated by polycomb group (PcG) proteins (box 4).

Box 4

Besides DNA methylation, CGIs can also be silenced by the polycomb group proteins (PcG). Originally identified in Drosophila as regulators of the Hox genes, PcG proteins are evolutionary conserved epigenetic modifiers with repressive functions. There are two PcG complexes in mammalian cells, polycomb-repressive complex 1 (PRC1) and PRC2. PRC2 mediates the tri-methylation of lysine 27 of histone 3 (H3K27me3) which in turn recruits the PRC1 complex and leads to the inhibition of transcriptional elongation and/or chromatin compaction [41,42]. In ES cells, CGIs silenced by the PRC2 complex possess the so-called "bivalent" domains. Those are marked by the active H3K4me3 modification, as well as by the repressive H3K27me3 mark [43]. Bivalent domains in ES cells are located within the CGIs of developmental genes which are in a poised state for expression. Upon differentiation they provide a quick "on-off" switch between active transcription and stable repression.

Similar to TET1, 5hmC was found to be enriched at TSS with intermediate- and high- CpG content and in gene bodies. What was somewhat surprising, the predicted association of TET1 and 5hmC with active state of promoters, was disproved by both studies. Moreover, TET1 contributed to transcriptional silencing via direct interaction with the SIN3a co-repressor complex [39].

An alternative approach for determining the position of 5hmC in the genome of mESCs was taken by Rao and colleagues. Two methods were developed [38] utilizing chemical conversion of 5hmC and subsequent ChIP followed by genome sequencing. Both the GLIB (glucosylation, periodate, oxidation, biotinylation) and CMS (cytosine 5-methylenesulphonate) techniques detected 5hmC at the promoters of genes with "poised" chromatin configuration [38].

Detailed re-assessment of the results generated by ChIP sequencing-based techniques [44,45], revealed that there was some degree of bias intrinsic to the

specific methodologies applied. For instance, anti-5hmC antibodies recognize more efficiently modification-dense regions [38], as well as CA repeats [40]. To overcome the limitations of affinity-based genome-wide profiling, methods for single-base-resolution mapping of 5hmC have been developed. These have attempted to provide more accurate and quantitative information regarding the precise position of 5hmC in the whole genome. Making use of different properties of modified cytosines, two groups independently designed modified bisulfate sequencing approaches for single-nucleotide-resolution mapping of both 5mC and 5hmC [46,47]. Booth et al. [42] made use of the fact that potassium perruthenate (KRuO4) specifically oxidizes 5hmC to 5fC which is subsequently deaminated under repeated bisulfite treatment. Consequently, in a DNA sample treated with KRuO4, 5hmC would be read as a thymine, whereas 5mC would be read as an unmodified cytosine. Oxidative-bisulfite sequencing (oxBS-Seq) is an elegant and simple approach allowing the concomitant detection of both 5hmC and 5mC without a prior enrichment step. The second method designed by Yu et al. [43] is called TAB-Seq, standing for TET-assisted bisulfite sequencing. Briefly, 5hmC is protected from the activity of the TET proteins, by blocking it with glucose using the β-glucosyltransferase. Next, all 5mCs are oxidized by mouse tet1 to 5caC which subsequently undergoes deamination during bisulfite treatment. Upon DNA sequencing these bases are read as thymines. The only remaining cytosine signals, generated after Tab-Seq, stem from the protected 5hmCs. This technique also does not rely on an enrichment step preceding the procedure. Tab-Seq has uncovered some new features of 5hmC such as its significant enrichment at distal functional regulatory elements, its distribution near transcription factor-binding sites, and the sequence bias and strand asymmetry associated with 5hmC sites.

TET2-binding sites on the genome were first reported by Chen et al. [48]. ChIP-sequencing was performed using anti-mouse Tet2 antibody in mESCs. The authors found that 47% of the 5hmC-positive genes were bound by Tet2. Most Tet2 targets were localized in intermediate- and high- CpG dense promoters, which were enriched for the active H3K4me3 mark. Unlike TET1, TET2 does not bind to the repressive SIN3a complex [39], but is prominently present in a complex with the O-linked beta-N-acetylglucosmaine transferase (OGT) protein [48]. TET2 facilitates the binding of OGT to chromatin and hence regulates histone 2B O-GlcNAcylation, a mark associated with active transcription. These data suggest that TET1 and TET2 may antagonize each other in the regulation of gene transcription in ESCs.

5fC and 5caC are continuous products of TET-mediated oxidation of 5hmC which are eventually excised by TDG to restore unmodified cytosines (see 2.2). The level of 5fC is about 10 fold lower than that of 5hmC in mESCs, and that of 5caC is yet 10 fold lower than that of 5fC. In fact, 5caC has so far only been detected in mESCs, but not in other tissues [31,49,50]. Mapping of 5fC and 5caC poses an even

greater challenge, given these cytosine species have a low abundance, and are expected to be much shorter lived, as compared to 5hmC. Raiber et al. [51] used a biotin-streptavidin based pulldown approach to determine the genome-wide distribution of 5fC in ESCs. Similar to 5hmC, the distribution of 5fC followed a similar pattern and was primarily found in euchromatic regions, including CGIs, exons and promoters. 5fC was also associated with active gene expression, with 5fC-rich CGI promoters corresponding to higher expression levels and presence of H3K4me3.

Intriguing findings regarding the putative function of the oxidized derivatives of 5mC have recently been published [52,53]. These studies attempted to identify specific readers of 5hmC, 5fC and 5caC in various cell types, including mESCs, neuronal progenitors and adult mouse brain tissue, through MS-based proteomic approaches. Several striking observations were reported. There was little overlap between the readers of the individual modifications in different cell types and tissues, indicating that these epigenetic marks may serve distinct cell-type-dependent roles. Few specific readers of 5hmC were detected, and those seemed to further confirm the role of DNA hydroxymethylation as an active DNA demethylation intermediate. Surprisingly however, a number of proteins, more than initially anticipated, were bound to 5fC and 5caC [52,53]. Apart from proteins involved in DNA damage response and base-excision repair pathways, 5fC recruited a considerable amount of transcriptional regulators and proteins, which are unlikely to be a part of the DNA demethylation pathway.

Further studies are required to clarify as to what extent, and if at all, the oxidized intermediates of 5mC function as stand-alone epigenetic marks.

2.4 TET2 in hematopoiesis: lessons from mice

Upon the discovery of *TET2* mutations in a multitude of hematological cancers, the question remained how the dysfunction of the protein deregulates hematopoiesis. Initially, Ko et al. demonstrated that genomic DNA obtained from bone marrow of patients carrying a *TET2* mutation displayed consistently lower 5hmC levels as compared to bone marrow samples of healthy individuals [54,55]. Overexpression of mutant TET2 proteins in cell lines further confirmed that TET2 defects perturb its catalytic activity [54,55]. In addition, it was shown that shRNA-mediated depletion of Tet2 in mouse hematopoietic precursors skewed differentiation towards the monocyte/macrophage lineage in cultures [54].

To address the issue *in vivo*, several groups engineered Tet2 knock out (KO) mouse models. The group of Bernard [56] used two independent approaches, a gene-trap model inactivating tet2 at the ES cell level, and mice bearing a conditional Tet2 allele. A marked reduction of 5hmC levels was detected in cells where Tet2 was absent. Analysis of the hematopoietic compartments of the affected mice revealed an amplification of the $Lin^{-}Sca^{+}Kit^{+}$ (LSK; hematopoietic stem cell)

pool. Compared to control animals, conditional Tet2 KO mice had increased absolute numbers of common myeloid progenitors (CMPs) and megakaryocyte-erythrocyte progenitors (MEPs), but the granulocyte-macrophage progenitor (GMP) population remained stable. When transplanted, total bone marrow or purified LSK cells from the Tet2-deficient mice, were both able to reconstitute all hematopoietic lineages in wild-type recipient mice. Over time, the contribution of Tet2-depleted cells increased in both the myeloid and the lymphoid lineages, and an amplification of the LSK compartment was observed in some recipient animals. Eventually, the gene-trap animals developed bona fide myeloid tumors displaying CMML-like disease features which were transplantable to secondary recipients. In line with these data, Li et al. [57] also showed that the bone marrow of tet2-null mice had an enriched LSK population and diminished 5hmC levels. Moreover, within 2-4 months these animals developed a CMML-resembling phenotype featuring persistent monocytosis, hepato- and splenomegaly and myeloproliferation. A significant percentage of Tet2+/- and Tet2-/- animals progressed to lethal myeloid malignancies, such as CMML, MPN-like myeloid leukemia, myeloid leukemia with maturation and MDS. In addition, Tet2-depleted cells demonstrated an obvious proliferative capacity in serial transplantation assays. Similar results were also reported by Ko et al. [58] whose conditional Tet2 KO mouse model further strengthened the evidence that bone marrow of mice lacking Tet2 showed a clear expansion of the hematopoietic stem cell pool. Tet2-/- cells had an enhanced ability to reconstitute hematopoiesis in vivo relative to tet2 wild type cells. Finally, Tet2 posed a constraint on cells from undergoing differentiation as shown by expression of lineage specific markers. All studies also uniformly proved that although expressed, Tet1 and Tet3 are unable to compensate for the loss of Tet2, neither by overexpression (mRNA), nor by maintaining normal 5hmC levels in the affected cells.

2.5 "Take your vitamins!"

Vitamin C (= L-ascorbic acid) is a water-soluble compound with essential nutritional properties. Among other animal species, primates have lost the capability to synthesize vitamin C, and require its intake through diet.

Although mainly popular for its antioxidant activity, vitamin C also serves as an electron donor to a number of enzymes, including DNA- and histone-modifying Fe (iron)-dependent oxidoreductases. Since most cell culture media do not contain vitamin C, the pivotal role of this compound in active DNA demethylation remained unacknowledged until a report appeared in 2010 showing that ascorbate can induce widespread demethylation in the epigenome of human ESCs [59]. Yet, it took a few more years to establish a clear link between the TET enzymes and vitamin C.

Several reports appeared in 2013 [60-62] demonstrating the direct role of vitamin C as an important co-factor of the TET proteins. As stated earlier in this introduction (2.1), TET

enzymes belong to the group of 2-oxoglutarate and Fe(II)-dependent dioxygenases (2OGFeDO). To convert 5mC into 5hmC, TETs utilize Fe(II) (also frequently denoted as Fe^{2+}), as a co-factor, and 2-oxoglutarate (2-OG) as a co-substrate. Shortly, molecular oxygen (O_2) is used for the oxidative decarboxylation of 2-OG into succinate and CO_2, leading to the formation of intermediate Fe(IV)-oxo species that enable hydroxylation of the methyl group of the substrate. If that occurs, Fe(IV) is reduced back to Fe(II) after release of the succinate. However, in the absence of a substrate succinate is released without reduction of iron to Fe(II), resulting in the generation of catalytically inactive oxidized iron species and enzyme inactivation. Minor et al. [60] hypothesized that to maintain the enzyme fully active, ascorbate is required, to serve as an electron donor for the reduction of oxidized iron species. They showed that the addition of vitamin C, time- and dose- dependently, enhanced the formation of 5hmC in mouse embryonic fibroblasts, without affecting the expression of the *Tet* genes.

Simultaneously, Blaschke et al. [61] made the serendipitous discovery that knockout serum replacement (KSR: a defined, serum-free formulation optimized to grow and maintain undifferentiated ESCs in culture) promoted TET-dependent demethylation, accompanied by an increase in 5hmC levels in mESCs and upregulated expression of germ-line genes. A small molecule screen identified vitamin C as the component responsible for this effect.

The relevance of vitamin C in the modulation of TET function and 5hmC levels should remain the subject of further exploration, more so in the light of the possibility to design vitamin C-based therapies for TET/5hmC-deficient cancers.

3. ACUTE MYELOID LEUKEMIA (AML)

3.1 Definition and classification of AML

Acute myeloid leukemia (AML) is a clinically and genetically heterogeneous clonal hematopoietic malignancy which arises in precursor cells produced in the bone marrow. As a consequence of cytogenetic and molecular lesions, the affected cells lose their ability to differentiate into mature blood cells of the myeloid lineage but have increased proliferative and self-renewal properties. The immature leukemic cells, i.e. blasts, expand and eventually completely overwhelm the bone marrow and blood, suppressing normal hematopoiesis. Due to the lack of mature erythrocytes, monocytes, granulocytes and thrombocytes, AML patients suffer from anemia, and an increased risk of infection and bleeding. AML is among the most common forms of acute leukemia in adults, and its incidence generally increases with age with ~70 years being the median age of onset. In elderly patients, the disease often presents as an evolution from MDS, whereas in younger patients *de novo* AML is more common. AML affects males more often than females.

In the past three decades, major advances have been made in understanding the genetic origins of AML and implementing this knowledge in the diagnosis,

prognosis and treatment of patients. The revised World Health Organization (WHO) classification of acute myeloid leukemias from 2008 accordingly recognizes subtypes of AML as unique disease entities based on specific recurrent genetic abnormalities that predict prognosis and response to therapy. Approximately 50-60% of the AML patients have an abnormal karyotype. Cytogenetic abnormalities have been extensively investigated and their prognostic value has been determined by several independent studies [63]. The most frequently observed karyotypic lesions and their impact on prognosis are listed in table 1.

Table 1. Prognostic classification of AML based on recurrent cytogenetic abnormalities

Favorable prognosis	Intermediate prognosis	Unfavorable prognosis
t(15;17)	Entities not classified as	abn(3q) excluding t(3;5)
t(8;21)	favorable or unfavorable	inv(3)/ t(3;3)
inv(16)		add(5q)/ del(5q)/ -5q
		add(7q)/ del(7q)/ -7q
		t(6;10)
		t(10;11)
		t(11q23) excluding t(9;11) and t(11;19)
		t(9;22)
		abn(17p)/ -17
		Monosomal karyotype*

Adapted from Grimwade et al. [63]
* at least two autosomal monosomies, or a single monosomy accompanied by at least one other structural chromosomal abnormality [64]

3.2 Common genetic mutations in AML

Nonetheless, AML patients with normal cytogenetics (CN-AML) still comprise the largest subgroup of AML patients (~45%) [65] which poses a particular complication in terms of prognostication and taking optimal treatment steps. Attempts to resolve this have resulted in the rapid identification of a myriad of new molecular markers in AML allowing for a more genetically-oriented risk stratification of patients, deeper understanding of the basic biology of AML, and the development of new targeted forms of therapy. Among the most frequently affected genes are *FLT3* (25%), *NPM1* (25-35%) *CEBPa* (10-15%), and *RAS* (10-15%) (Table 2).

The fms-like tyrosine kinase 3 (FLT3) is a receptor tyrosine kinase which plays an important role in governing normal hematopoiesis. *Flt-3* alterations include internal tandem duplication (ITD) of the juxtamembrane domain, as well as mutations in the tyrosine kinase domain (TKD) of the protein. AML patients who carry a *FLT3*-ITD

Table 2. Prognostic significance of recurrent genetic mutations in AML

Gene	Incidence	Prognostic impact	Gene function
NPM1	25-35%	Favorable	Nuclear import, apoptosis
FLT3-ITD	20-25%	Unfavorable	Growth factor receptor
DNMT3a	20-23%	Unfavorable	Epigenetic regulation
MLL	10-15%	Unfavorable	Epigenetic regulation
N/K-RAS	10-15%	Intermediate	Signal transduction
CEBPa	10-15%	Favorable (bi-allelic mutations)	Transcription factor
CBFB-MYH11 Inv(16)(p13q22)	10%	Favorable	Transcription factor
AML1-ETO t(8;21)(q22;q22)	10%	Favorable	Transcription factor
TET2	8-13%	Unfavorable	Epigenetic regulation
IDH2	6-10%	Unclear	Metabolism
PML-RAR t(15;17)(q22;q21)	5-10%	Favorable	Transcription factor
ASXL1	5-10%	Unfavorable	Epigenetic regulation
RUNX1	5-10%	Unfavorable	Transcription factor
FLT3-TKD	5-7%	Intermediate	Growth factor receptor
WT1	3-10%	Unclear/unfavorable	Transcription factor
IDH1	3-9%	Unclear	Metabolism
PTPN11	3-5%	Intermediate	Signal transduction
C-KIT	2-6% (high in CBF leukemia)	Unfavorable in CBF leukemia	Growth factor receptor
CBL	<2% (16% in inv16)	Unclear	Signal transduction

have a worse prognosis [66,67] and the significance of *FLT3*-ITD as an independent poor prognostic factor has been established by multiple studies [67,68]. Unlike ITD, *FLT3*-TKD mutations are rarer and do not appear to have an adverse effect on prognosis [69]. The flt-3 pathway is an obvious target for the development of small molecule inhibitors that block the constitutive activation of the mutant protein and some are already in clinical and preclinical development [70]. For instance, quizartinib (AC-220) and midostaurin (PKC412) have shown promise in certain AML subgroups in phase II/III clinical trials, and sorafenib has been reported to show a prominent activity against *FLT3*-ITD positive AML [71].

Nucleophosmin 1 (*NPM1*) mutations, despite being among the most frequent genetic aberrations in AML (25-35% in AML; 50-60% within the subgroup of cytogenetically normal CN-AML), were discovered no earlier than 2005 [72]. Albeit rare, *NPM1* was already known to be a translocation partner of several proteins, among which *RAR*a [73] in acute promyelocytic leukemia (APL). This prompted further research which led to the key publication by Falini and colleagues [72], describing the aberrant cytoplasmic, instead of nuclear, localization of NPM1 in patients with

AML caused by mutations affecting exon 12 of the gene. The presence of *NPM1* mutations is generally considered a favorable prognostic factor, in the absence of *FLT3*-ITD [74].

CCAAT/enhancer-binding protein alpha (CEBPa) is a transcription factor which plays a critical role in myeloid development. Mutations of *CEBPa* are associated with significantly better outcome in CN-AMLs, only in the case of biallelic aberrations [75]. Like *NPM1*, the group of *CEBPa* mutated AML patients, is now defined as a provisional entity with unique genetic characteristics in the refined version of the 2008 WHO classification of myeloid neoplasms and acute leukemias [76].

Activating mutations in the *N-* and *K-RAS* proto-oncogenes occur in 10-15% of AML and do not appear to significantly affect survival [65], although one study has shown that *RAS* positive leukemic cells may be more sensitive to high-dose cytarabine treatment [77].

4. EPIGENETIC DEREGULATION IN AML

For over a decade, the classical two-hit model of leukemogenesis dominated the hemato-oncology field [78]. Essentially, it stated that at least two genetic events affecting different classes of proteins were required to drive malignant transformation of the hematopoietic progenitors. Class I genes, affecting proliferation, provide the transformed clone with growth advantage, whereas class II genes, involved in differentiation, alter the execution of normal myelopoiesis. Although useful, this model is likely oversimplified and has been challenged over the years through the discovery of novel classes of mutations in myeloid neoplasia [79], such as mutations affecting RNA splicing, which have recently been discovered in MDS [80]. Another distinct group are genetic defects perturbing epigenetic regulation, either at the DNA level (DNA (de)methylation), or at the histone level (various histone modifiers).

4.1 DNA-methyltransferases (DNMTs) in AML

Aberrant DNA methylation profiles are a hallmark of leukemia. Promoter hypermethylation inactivates a number of tumor suppressor genes involved in a variety of cellular processes and signaling pathways to promote cancer proliferation. Of note, methylation state is not only useful as a biomarker, but it has also been associated with prognosis and treatment response. For instance, hypermethylation of the Wilms tumour (WT)1 gene in AML has been shown to correlate with a chemotherapy resistant phenotype [81]. Furthermore, AML patients display heterogeneous patterns of methylation of the cyclin-dependent kinase (CDKN)2B [82]. The methylation signature of CDKN2B marks disease progression [83]. In 2001 Mizuno et al. [84] showed that *DNMT1*, *DNMT3a* and *DNMT3b* are all overexpressed in AML patients, as well as in cells of patients in acute phase chronic

myeloid leukemia (CML). *DNMT1* and *DNMT3b* are specifically overexpressed in AMLs with CDKN2B hypermethylation suggesting involvement of these proteins in the aberrant repression of the cyclin kinase. In CML, overexpression of the *DNMT*s seems to correlate with disease evolution, as normal levels of the *DNMT*s are present in chronic phase CML, which increase upon progression to blast crisis CML.

Interestingly, following deep sequencing of the genomes of two AML patients with normal karyotype, the group of Ley [85] discovered a frameshift mutation in the *DNMT3a* gene. Sanger sequencing of a cohort of 281 AML patients identified *DNMT3a* defects in 22,1% of the cases. The most commonly affected amino acid was R882, although nonsense, missense, frameshift and splices-site aberrations were found in other portions of the protein as well. Moreover, *DNMT3a* mutations represented an independent poor prognostic factor in AML; they were significantly correlated with *FLT3* and *NPM1*, and cytogenetically normal AML, as well as AML- M4 and AML-M5 FAB subtype. DNA methylation was measured by liquid chromatography (LC) - tandem mass spectrometry (MS) but revealed no differences between wild type and DNMT3a mutant cells. MeDIP analysis showed decreased DNA methylation at 182 loci but there was no consistent correlation with the expression of the respective neighboring genes.

Shortly after, Yan et al. [86] reported the identification of somatic mutations in *DNMT3a* upon exome sequencing of acute monocytic leukemia (AML-M5) in 20,5% of the cases. Additional screening of other AML FAB subtypes affirmed the presence of *DNMT3a* mutations in AML-M4 as well. Patients with aberrant *DNMT3a* had inferior overall survival. *In vitro* data suggested reduced enzymatic activity of the mutant DNMT3a protein and reduced affinity to histone H3. Of note, DNA methylation and gene expression patterns were altered (e.g HOXB genes) in cases bearing a mutation as compared to wild type counterparts.

At present, the precise mechanism(s) through which *DNMT3a* mutations contribute to leukemogenesis remains unclear. Most aberrations found in patients are heterozygous suggesting that haploinsufficiency is capable of driving myeloid transformation. However, *in vitro* studies using purified protein do not properly reflect the *in vivo* situation where the wild type allele is still present. Studies assessing the effect of *DNMT3a* mutations on genome 5mC levels are conflicting [85-87]. Several scenarios seem plausible. *DNMT3a* mutations may inhibit the unaffected allele in a dominant negative fashion, since DNMT3a is known to form oligomers. Alternatively, they may disturb interactions with other proteins, such as DNMT3L. DNMT3L is a shorter form which lacks methyltransferase activity, but increases the binding of DNMT3a to certain loci [88,89].

Surprisingly, conditional disruption of either DNMT3a or DNMT3b in mouse CD34- LSK cells did not affect hematopoiesis. Nevertheless, conditional depletion of both DNMT3a and DNMT3b did. The de novo methyltransferases were essential for self-renewal, but not for differentiation, in hematopoietic

cells. Cells lacking both DNMT3a and DNMT3b retained myeloid and lymphoid lineage differentiation potential but were incapable of long-term reconstitution in transplantation assays [90]. More recently, Goodell and colleagues showed that conditional ablation of DNMT3a in the hematopoietic system impaired HSC differentiation over serial transplantations and resulted in an expansion of the stem cell population in the bone marrow [91]. In this model, both hypomethylation and hypermethylation of CpG islands were found. DNMT3a-null animals showed an upregulation of multipotency genes whereas transcription factors were downregulated.

Provided the role of DNMT3a mutations in AML, as well as the combined contribution of both de novo methyltransferases in leukemia mouse models, it came as no surprise when a role for DNMT3b in AML was established. Sequence analysis showed the absence of acquired mutations in *DNMT3b.* However, overexpression of *DNMT3b* in 192 *de novo* AML patients was an independent indicator of poor survival [92,93].

4.2 The metabolic enzymes connection

In 1924 Otto Warburg postulated his famous hypothesis, stating that metabolic insults are the fundamental cause of cancer. Nowadays, deregulation of multiple oncogenes and tumor suppressors is known to alter metabolism and induce aerobic glycolysis [94]. Little did Warburg know at the time, that mutations in metabolic enzymes themselves would turn out to be more multifaceted than expected.

Two key publications appeared in 2009, showing the frequent involvement of the isocitrate dehydrogenase 1 (IDH1) and IDH2 in cancer. Yan et al. [95] reported hotspot mutations affecting *IDH1* in 70% of glioma patients and to a lesser extent *IDH2*. Mardis et al. [96] found *IDH1* mutations in 8,5% of AML patients. The overall prevalence of *IDH1* and *IDH2* mutations in AML was ~20%. In contrast to gliomas where *IDH1* mutations were more frequent, in AML *IDH2* mutations were more common suggesting a potentially different oncogenic effect of IDH in gliomas versus AML [97].

The IDH1 and IDH2 are NADP(+)-dependent metabolic enzymes which catalyze the oxidative decarboxylation of isocitrate into alpha-ketoglutarate (α-KG, also 2-oxoglutarate 2-OG). IDH1 is located in the cytoplasm and peroxisomes, whereas IDH2 is a mitochondrial component of the Krebs cycle (TCA cycle). Initially, thought to have a dominant-negative effect, *IDH* mutations turned out to have a neomorphic activity instead. The defective proteins gain the novel function of converting α-KG into the R(-) enantiomer of 2-hydroxyglutarate (further on referred to as simply 2-HG). The build-up of the onco-metabolite 2-HG was not restricted to cell extracts but was also found to rapidly accumulate in the medium of cells expressing the mutant forms [98]. Elevated 2-HG has been found both in cells and serum of AML patients [97,99,100].

The most common *IDH1* mutation affects amino acid 132, an arginine residue (R132) in the active site of the enzyme. The mitochondrial homolog, IDH2, bears two hotspot substitutions affecting the analogous R172 residue, and R140 which has also been predicted to bind and orient α-KG in the active site [81]. *IDH* aberrations are always heterozygous retaining one healthy allele. Since IDH proteins exist as homodimers, the mutant allele promotes NADPH-dependent reduction of α-KG provided by the wild type allele to generate 2-HG [98]. Elevated 2-HG however, is not exclusive to IDH transformed tumors. Toxic accumulation of the R(-), as well as the S(-) stereoisomers of 2-HG have also been observed in patients carrying rare germ-line mutations in the D(R)- and L(S) -2 hydroxyglatarate dehydrogenases, D- and L- 2HGDH respectively. Interestingly, some L-2HG aciduria patients develop brain tumors. Although brain cancer has not been reported in D-2HG aciduria individuals, D-2HG aciduria follows a more severe clinical course and patients likely succumb before gliomas can form [101-103]. This explanation may also be a valid reason why hematological disorders have never been reported in patients suffering from this inborn error of metabolism. Of note, in previously unexplained D-2HG aciduria patients lacking mutations in the D-2HGDH enzyme, *IDH2* mutations appeared to be causative [104].

In 2010 Figueroa et al. [105] showed that the dominant feature of bone marrow CD34+ AML cells with mutated *IDH1/2* was hypermethylation of a subset of promoters. The DNA hypermethylation signature was similar between *IDH1* and *IDH2* mutant AMLs suggesting a shared pathogenic mechanism in the epigenetic deregulation of gene expression. Ectopic expression of IDH1 R132 and IDH2 R172 mutants in HEK293T cells led to a significant increase in 5mC levels. As the TET proteins utilize α-KG (= 2-OG), Fe(II) and molecular oxygen as co-factors in the generation of 5-hmC, the potential link between *IDH* mutations and *TET2* mutations soon became evident (Fig. 5). Moreover, *TET2* and *IDH* aberrations mutually exclusive, yet again alluding that alterations in these proteins perturb a common biological pathway [105]. Co-transfection of TET2 with mutant IDH1/2, but not the wild type proteins, inhibited the formation of 5hmC in HEK293T cells. In addition, TET2 and IDH1/2 hypermethylation signatures in AML overlapped.

Later, a conditional knock-in (KI) IDH1 (R132H) mouse model was generated by Sasaki et al. [106]. The mutant form was inserted either in all hematopoietic cells or in cells of the myeloid lineage specifically. Analogous to the phenotype of TET2 KO mice (see section 2.4), the bone marrow of IDH1 KI animals was eventually overgrown by immature hematopoietic progenitors (LSK cells and lineage-restricted progenitors, LRK cells). Expression arrays performed on sorted LSK cells from KI and control animals, showed alterations in a set of genes related to cellular growth and proliferation. A significant amount of CpG sites was aberrantly methylated demonstrating the hypermethylator activity of the IDH1 R132 mutant *in vivo*. The WNT, Notch, and TGF-β pathways which are all targeted by promoter and intragenic hypermethylation in IDH1/2 mutant AML cells were also implicated in the IDH1 KI mouse model.

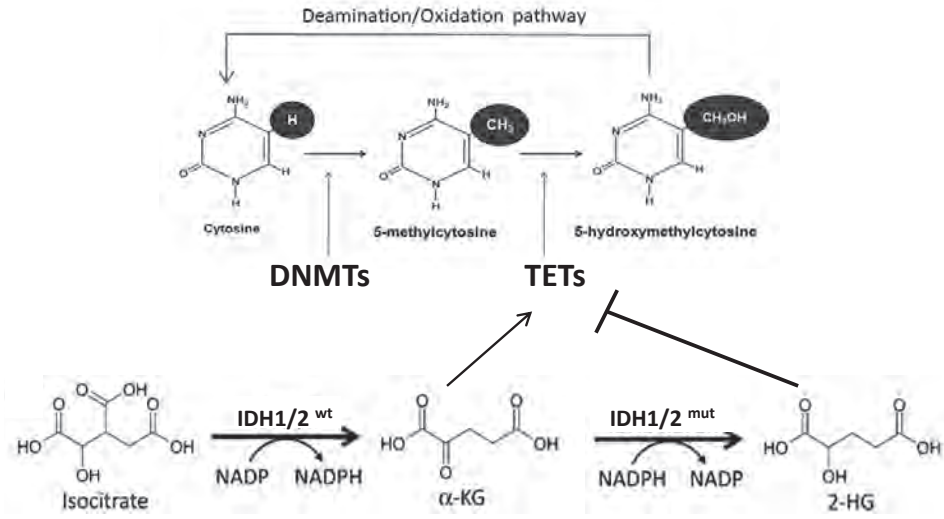


Figure 5. *2-HG is a competitive inhibitor of the TET proteins. IDH1 and IDH2 convert isocitrate into α-KG (= 2-OG), either in the cytosol (IDH1) or in the mitochondria during the Krebs cycle (IDH2). Whereas the wild type IDH proteins supply α-KG to the TET enzymes, the mutants gain the novel function of further processing α-KG into 2-hydroxyglutarate (2HG). 2-HG accumulates in the affected cells, as well as their surrounding environment, and competes with α-KG for binding at the same active sites*

Notwithstanding, the proof-of-principle studies linking the production of 2-HG with the functional inhibition of the TET enzymes, placed *IDH* mutations in a much broader context.

TET proteins, as mentioned in section 2.1, belong to the superfamily of 2-oxoglutarate and Fe(II)-dependent dioxygenases (note 2).

Note 2

2-oxoglutarate and Fe(II)-dependent dioxygenases are a class of enzymes that are conserved in eukaryotes and bacteria. They catalyze a variety of reactions typically involving the oxidation of an organic substrate using a dioxygen molecule, mostly using ferrous iron Fe(II) as the active site co-factor and 2-OG as a co-substrate which is decarboxylated to succinate and CO2.

Mammalian cells express over 60 dioxygenases that utilize α-KG (= 2-OG) as a co-substrate among which the TET enzymes, and JmjC-domain containing histone demethylases (JHDMs) [107-109]. The JmjC demethylases demethylate mono-, di-, and tri- methylated lysine residues on histone H3 [110,111]. 2-HG produced by IDH mutants does not only inhibit the TETs but also multiple histone demethylases *in vitro* and *in vivo* [112]. Overexpression of IDH1 R132 in a glioma cell line led to the visible

increase in H3K4me1, H3K27me2, H3K4me3, H3K9me2, and H3K79me2 levels which was reversible upon the addition of cell permeable α-KG. Furthermore, a panel of human glioma samples with or without IDH1 mutation was tested. H3K79me2 was significantly elevated in tumors carrying an *IDH1* defect. Later, Lu et al. [113] investigated the effects of IDH mutants and 2-HG on differentiation using 3T3-L1 cells as a model. Upon stimulation with a differentiation cocktail, immortalized murine 3T3-L1 cells undergo extensive chromatin remodelling leading to their maturation into adipocytes. Transduction of these cells with either *IDH2* R172 or *IDH2* R140 mutant resulted in a block in differentiation, which was not the case for cells transduced with wild type *IDH2*. Furthermore, cell permeable 2-HG alone was able to induce the same phenotype. Gene expression analysis of cells transduced with *IDH* or treated with 2-HG showed a profound defect in the expression of transcription factors essential for adipocyte-lineage commitment. Biochemical data revealed an increase in H3K27me3 and H3K9me3. In addition, the authors tested normal human astrocytes transduced with *IDH1* R132 and demonstrated that long-term cultures displayed a profound elevation of H3K9me3 and CpG DNA methylation.

Taken together, a body of evidence has been compiled examining the wide-range defects induced by *IDH* mutations in cancer cells. Given that 2-HG not only accumulates in cells but also in the surrounding environment, an important question still remains to be addressed. It would be of tremendous interest to determine whether 2-HG excreted by the leukemic cells is able to penetrate nearby healthy cells and tissues thus modifying their transcriptional program. Notably, both in MDS and AML local changes in the hematopoietic niche are present and complex interactions between the transformed cells and their microenvironment exist [114].

4.3 Histone modifiers in AML

Epigenetic deregulation due to mutations in histone modifiers is not a rare event. Mixed-lineage Leukemia (*MLL*)-rearranged leukemias bearing chromosomal translocations involving chromosome 11q23 represent ~10% of the adult AML cases. There are over 60 known fusion partners of MLL (among which TET1, see section 2.1) which correlate with disease phenotype and prognosis albeit the presence of *MLL* defects generally confers poor prognosis [115]. MLL belongs to the trithorax group (TrxG) of chromatin modifiers and it regulates the expression of *Hox* genes which play a critical role in hematopoiesis [116,117]. The MLL protein contains an N-terminal DNA binding domain, and a C-terminal SET domain with H3K4 methyltransferase activity [118,119]. Biochemical studies have identified MLL as a member of a large multiprotein complex that contains other chromatin remodelling proteins, such as histone deacetylases (HDACs) and components of the Swi/Snf chromatin-remodelling complex [103]. All translocations involving *MLL* appear to produce a fusion protein possessing the N-terminal portion of MLL fused in frame to the C-terminus of the fusion partner[120]. Essentially, *MLL* retains its DNA binding

capability but is bereft of its methyltransferase activity. Pathways targeted by *MLL* fusions include DNA methylation, histone acetylation and histone methylation [115].

A member of the PRC2 complex (see box 4, section 2.3) has recently been implicated in myeloid malignancies. The Enhancer of Zeste Homologue 2 (EZH2) is an H3K27 methyltransferase. Mutations and overexpression of *EZH2* have both been reported in AML. Alterations of wild type EZH2 expression in MDS/AML have in addition been associated with inferior outcome [121]. Intriguingly, *EZH2* harbors both activating and loss-of-function mutations suggesting its transforming contribution as both an oncogene and a tumor suppressor. Initially, the gain-of-function Y641 mutation was detected in patients with germinal-centre diffuse large-B cell lymphoma [122] and it was further demonstrated that indeed both di- and tri- methylation of H3K27 were increased [123]. Later, recurrent nonsense, missense and frameshift aberrations were identified in patients with MDS [124], CMML and primary myelofibrosis (PMF) [125]. *In-vitro* data have shown that those are likely inactivating[126]. Mutations of *EZH2* in AML are in fact rare [127,128].

Although infrequent mutations and/or chromosomal rearrangements involving other histone (de)methylases, such as *NSD1* (H3K36 methyltransferase), *JARID1A* (KDM5a, lysine demethylase) and *UTX* (KDM6b, H3K27 demethylase) have also been reported [129].

4.4 Epigenetic therapies in MDS and AML

As more and more whole AML genomes are being sequenced, a few progressive concepts have emerged: 1) The list of the recurrent "culprits" is now (almost) complete; 2) Leukemia arises as a multistep process following the evolution of a single transformed clone which expands, but despite what was initially believed, this clone carries only a few driver mutations and a snapshot of multiple "by-standers". This essentially means that mutations are accumulated in the hematopoietic cells of healthy individuals as a result of aging. The AML genome thus already contains a number of mutations that were present before the initiating pathogenic event. The transformed clone(s) accumulates additional insults yielding subclones that can contribute to disease progression and/or relapse [130]. Pappaemmanuil et al. even suggest the possibility of genetic "predestination", meaning that early genetic insults set cancer clones on a certain path of clonal evolution by limiting the repertoire of cooperating lesions. Moreover, the number of driver mutations correlates with the outcome of patients [131]; 3) A necessity arises to translate our knowledge of the genetic background of AML into targeted forms of therapy.

The reversible nature of epigenetic marks makes them a very promising target for the development of novel therapies, especially within the group of elderly patients, where aggressive treatment regimens and hematopoietic stem cell transplantation (HSCT) are not advisory and treatment options remain limited.

Several drugs aimed at both DNA methylation and histone modifications have been approved for use in patients, and many are currently undergoing assessment stages in clinical trials. The first three epigenetic therapeutics included two DNMT inhibitors (DNMTi), namely azacitidine (5-aza = Vidaza) and its deoxy- derivative decitabine (5-aza-2' deoxycitidine), as well as the histone deacetylase (HDAC) inhibitor (vorinostat). Both 5-aza and decitabine are analogues of the naturally occurring pyrimidine cytidine. They are believed to exert their anti-neoplastic effect through two mechanisms: cytotoxicity through incorporation into DNA and/or RNA, and DNA hypomethylation through covalent binding to the DNMT1 protein resulting in its inhibition [132,133]. Whereas 5-aza incorporates to a larger extent into RNA rather than DNA, decitabine can only incorporate into DNA. Although the precise mechanism of action of DNMT and HDAC inhibitors remains elusive, they have shown promising clinical response in high-risk MDS and AML patients [134]. Due to their lower toxicity, as compared to traditional chemotherapeutics, DNA hypomethylating agents are often a preferred and safer alternative for frail patients.

Considering the implications of *TET2* mutations in myeloid malignancies on DNA methylation, it has been hypothesized that treatment of patients carrying *TET2* aberrations with hypomethylating agents may be beneficial. A small retrospective study from Itzykson et al. [135] conducted in 86 patients with high-risk MDS and low blast count AML, showed that *TET2* mutated cases had an 82% improved response rate versus 45% in *TET2* wild type patients. However, response duration and overall survival were comparable between the two groups. In contrast, another study reported that *TET2* alterations rather predicted decreased responsiveness to hypomethylating agents in MDS [136]. Sekeres et al. [137] assessed the efficacy of a treatment regimen combining lenalidomide and azacitidine in a cohort of 36 MDS patients. The authors claimed that patients with mutations affecting the methylation pathway achieved complete remission (CR) more often than patients without these mutations. However, such a claim is an overstatement, considering only 4 of the 36 patients carried a mutation in a gene involved in DNA (de)methylation (TET2, DNMT3a, IDH1/2). The small numbers of patients and the incomplete characterization of patient characteristics must be taken into account regarding the aforementioned reports. Larger studies with well-defined parameters and homogeneous patient groups will prove to be much more informative in assessing the potential benefits of DNA hypomethylating agents in patients carrying not only *TET2*, but also *IDH1/2* and *DNMT3a* mutations.

Hypermethylation of various histone lysine residues as a consequence of mutations in histone methyltransferases, histone demethylases and *IDH1/2*, have prompted attempts to target histone methylation in myeloid cancers. Among the most auspicious therapeutic compounds currently undergoing clinical assessment is EPZ-5676, a DOT1L inhibitor targeting *MLL*-rearranged leukemias. The four most frequent *MLL* translocations (*MLL-AF4*, *MLL-AF9*, *MLL-AF10* and *MLL-ENL*) result

in the recruitment of the disruptor of telomeric silencing 1 (DOT1L) to the fusion protein and consequentially an increase in H3K79 methyltransferase activity. EPZ-5676 is the first histone methyltransferase (HMT) inhibitor to enter clinical development. A phase I trial has been initiated in September 2012 (Epizyme).

EZH2 mutations result in gain- or loss-of-function affecting H3K27 methylation. Moreover, the H3K27 demethylase, UTX, has also been implicated in some of the same disorders as EZH2, providing a strong rationale for the development of targeted therapies aiming at H3K27 methylation. In fact, compounds able to pharmacologically inhibit EZH2 are already being explored [138-140].

OUTLINE OF THIS THESIS

The main objective of this thesis was to investigate in-depth the clinical significance, as well as the biological (dys)function of the novel epigenetic regulator TET2. As 30% of MDS patients eventually progress to AML, the first aim of this project was to determine the significance of TET2 mutations on survival of AML patients, and how these abrogate the function of the TET2 protein. Chapter 2 describes the study of TET2 mutations in 357 *de novo* AML patients who were enrolled in the EORTC/GIMEMA 06991 clinical trial. We show that TET2 mutation is an independent factor of poor prognosis in AML. Aberrant TET2 expression may also negatively affect overall survival of AML patients. Furthermore, we demonstrate that TET2 defects result in a loss of its DNA hydroxymethylation activity. Finally, we hypothesize that mutations affecting DNA hydroxymethylation (TET2, IDH1, IDH2) and DNA methylation (DNMT3a) likely cooperate during leukemogensis, based on their significant co- occurrence in patients.

In Chapter 3, the levels of 5-hydroxymethylcytosine in the same cohort of patients were determined (whenever possible), and correlated to clinical outcome. A highly quantitative HPLC-tandem MS method was developed and employed to measure 5hmC in genomic DNA. *TET2* and *IDH1/2* mutations clearly caused a significant decrease in DNA hydroxymethylation. Of note, 5hmC levels varied within a wide range. To exclude the effect of *TET2* and *IDH1/2* defects on overall survival, the study further focused on AML patients who were wild type for these aberrations. Unlike patients with intermediate levels, patients with high 5hmC levels had an inferior prognosis. In addition, high 5hmC content was strongly associated with the presence of *AML1-ETO* and *MLL1* translocations.

Chapter 4 of this manuscript poses the question whether 5hmC, like its preceding mark, 5mC, is maintained faithfully during cell division. In-depth analysis of TET (TET1, TET2, TET3) mRNA and protein expression in G0/G1, S (early, mid, late) and G2/M phases of the cell cycle, showed no significant changes. However, HLPC-MS-based measurements indicated that unlike 5mC levels, which remained stable due to the activity of DNMT1, 5hmC levels declined in late S and G2/M phases of the

cell cycle. This means that 5hmC is not copied immediately during DNA replication. We hypothesize that 5hmC, unlike 5mC, is not a stand-alone epigenetic mark, but a mere intermediate along the DNA demethylation pathway, and its formation is a consequence of pre-exisiting conditions, such as DNA methylation and histone modifications, rather than faithful maintenance during cell cycle.

REFERENCES

1. Langemeijer SMC, Kuiper RP, Berends M, et al: Acquired mutations in TET2 are common in myelodysplastic syndromes. Nature Genetics 41:838-U102, 2009
2. Delhommeau F, Dupont S, Della Valle V, et al: Mutation in TET2 in Myeloid Cancers. New England Journal of Medicine 360:2289-2301, 2009
3. Viguie F, Aboura A, Bouscary D, et al: Common 4q24 deletion in four cases of hematopoietic malignancy: early stem cell involvement? Leukemia 19:1411-1415, 2005
4. Vainchenker W, Delhommeau F, Constantinescu SN, et al: New mutations and pathogenesis of myeloproliferative neoplasms. Blood 118:1723-35, 2011
5. Ceredig R, Rolink AG, Brown G: Models of haematopoiesis: seeing the wood for the trees. Nature Reviews Immunology 9:293-300, 2009
6. Tefferi A, Levine RL, Lim KH, et al: Frequent TET2 mutations in systemic mastocytosis: clinical, KITD816V and FIP1L1-PDGFRA correlates. Leukemia 23:900-904, 2009
7. Jankowska AM, Szpurka H, Tiu RV, et al: TET2 gene harbors mutations associated with myeloid malignancies. Leukemia Research 33:S84-S84, 2009
8. Muramatsu H, Makishima H, Jankowska AM, et al: Mutations of an E3 ubiquitin ligase c-Cbl but not TET2 mutations are pathogenic in juvenile myelomonocytic leukemia. Blood 115:1969-1975, 2010
9. Langemeijer SMC, Jansen JH, Hooijer J, et al: TET2 mutations in childhood leukemia. Leukemia 25:189-192, 2011
10. Ismael O, Shimada A, Hama A, et al: De novo childhood myelodysplastic/ myeloproliferative disease with unique molecular characteristics. British Journal of Haematology 158:129-137, 2012
11. Busque L, Patel JP, Figueroa ME, et al: Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nature Genetics 44:1179-1181, 2012
12. Lorsbach RB, Moore J, Mathew S, et al: TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 17:637-641, 2003
13. Iyer LM, Tahiliani M, Rao A, et al: Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle 8:1698-1710, 2009
14. Tahiliani M, Koh KP, Shen YH, et al: Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324:930-935, 2009
15. Kriaucionis S, Heintz N: The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929-30, 2009
16. Wyatt GR, Cohen SS: The Bases of the Nucleic Acids of Some Bacterial and Animal Viruses - the Occurrence of 5-Hydroxymethylcytosine. Biochemical Journal 55:774-782, 1953
17. Qiu J: Epigenetics: Unfinished symphony. Nature 441:143-145, 2006
18. Deaton AM, Bird A: CpG islands and the regulation of transcription. Genes Dev 25:1010-22, 2011

19. Stein R, Gruenbaum Y, Pollack Y, et al: Clonal inheritance of the pattern of DNA methylation in mouse cells. Proc Natl Acad Sci U S A 79:61-5, 1982
20. Payer B, Lee JT: X Chromosome Dosage Compensation: How Mammals Keep the Balance. Annual Review of Genetics 42:733-772, 2008
21. Ooi SKT, Bestor TH: The colorful history of active DNA demethylation. Cell 133:1145-1148, 2008
22. Jin B, Li Y, Robertson KD: DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes Cancer 2:607-17, 2011
23. Gu TP, Guo F, Yang H, et al: The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477:606-10, 2011
24. Wossidlo M, Nakamura T, Lepikhov K, et al: 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun 2:241, 2011
25. Seisenberger S, Andrews S, Krueger F, et al: The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell 48:849-62, 2012
26. Popp C, Dean W, Feng SH, et al: Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature 463:1101-U126, 2010
27. Bhutani N, Brady JJ, Damian M, et al: Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463:1042-U57, 2010
28. Conticello SG, Langlois MA, Neuberger MS: Insights into DNA deaminases. Nature Structural & Molecular Biology 14:7-9, 2007
29. Di Nola JM, Neuberger MS: Molecular mechanisms of antibody somatic hypermutation. Annual Review of Biochemistry 76:1-22, 2007
30. Guo JU, Su YJ, Zhong C, et al: Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain. Cell 145:423-434, 2011
31. Ito S, Shen L, Dai Q, et al: Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science 333:1300-1303, 2011
32. He YF, Li BZ, Li Z, et al: Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science 333:1303-1307, 2011
33. Maiti A, Drohat AC: Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine POTENTIAL IMPLICATIONS FOR ACTIVE DEMETHYLATION OF CpG SITES. Journal of Biological Chemistry 286:35334-35338, 2011
34. Nabel CS, Manning SA, Kohli RM: The Curious Chemical Biology of Cytosine: Deamination, Methylation, and Oxidation as Modulators of Genomic Potential. Acs Chemical Biology 7:20-30, 2012
35. Rangam G, Schmitz KM, Cobb AJA, et al: AID Enzymatic Activity Is Inversely Proportional to the Size of Cytosine C5 Orbital Cloud. Plos One 7, 2012
36. Branco MR, Ficz G, Reik W: Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nature Reviews Genetics 13:7-13, 2012
37. Huang Y, Pastor WA, Shen YH, et al: The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing. Plos One 5, 2010
38. Pastor WA, Pape UJ, Huang Y, et al: Genome-widemapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473:394-397, 2011
39. Williams K, Christensen J, Pedersen MT, et al: TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473:343-U472, 2011
40. Wu H, D'Alessio AC, Ito S, et al: Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 473:389-U578, 2011

41. Stock JK, Giadrossi S, Casanova M, et al: Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nature Cell Biology 9:1428-U174, 2007
42. Eskeland R, Leeb M, Grimes GR, et al: Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. Molecular Cell 38:452-464, 2010
43. Mikkelsen TS, Ku MC, Jaffe DB, et al: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553-U2, 2007
44. Matarese F, Pau ECD, Stunnenberg HG: 5-Hydroxymethylcytosine: a new kid on the epigenetic block? Molecular Systems Biology 7, 2011
45. Song CX, Yi CQ, He C: Mapping recently identified nucleotide variants in the genome and transcriptome. Nature Biotechnology 30:1107-1116, 2012
46. Yu M, Hon GC, Szulwach KE, et al: Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. Cell 149:1368-1380, 2012
47. Booth MJ, Branco MR, Ficz G, et al: Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution. Science 336:934-937, 2012
48. Chen Q, Chen YB, Bian CJ, et al: TET2 promotes histone O-GlcNAcylation during gene transcription. Nature 493:561-+, 2013
49. Globisch D, Pearson D, Hienzsch A, et al: Systems-Based Analysis of Modified tRNA Bases. Angewandte Chemie-International Edition 50:9739-9742, 2011
50. Pfaffeneder T, Hackner B, Truss M, et al: The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew Chem Int Ed Engl 50:7008-12, 2011
51. Raiber EA, Beraldi D, Ficz G, et al: Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. Genome Biol 13:R69, 2012
52. Spruijt CG, Gnerlich F, Smits AH, et al: Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. Cell 152:1146-59, 2013
53. Iurlaro M, Ficz G, Oxley D, et al: A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome Biol 14:R119, 2013
54. Ko M, Huang Y, Jankowska AM, et al: Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature 468:839-43, 2010
55. Kroeze LI, Aslanyan MG, van Rooij A, et al: Characterization of acute myeloid leukemia based on levels of global hydroxymethylation. Blood 124:1110-8, 2014
56. Quivoron C, Couronne L, Della Valle V, et al: TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell 20:25-38, 2011
57. Li Z, Cai XQ, Cai CL, et al: Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood 118:4509-4518, 2011
58. Ko M, Bandukwala HS, An J, et al: Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proceedings of the National Academy of Sciences of the United States of America 108:14566-14571, 2011
59. Chung TL, Brena RM, Kolle G, et al: Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. Stem Cells 28:1848-55, 2010
60. Minor EA, Court BL, Young JI, et al: Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of

5-hydroxymethylcytosine. J Biol Chem 288:13669-74, 2013

61. Blaschke K, Ebata KT, Karimi MM, et al: Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500:222-6, 2013
62. Chen J, Guo L, Zhang L, et al: Vitamin C modulates TET1 function during somatic cell reprogramming. Nat Genet 45:1504-9, 2013
63. Grimwade D, Hills RK, Moorman AV, et al: Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. Blood 116:354-365, 2010
64. Breems DA, Lowenberg B: Acute myeloid leukemia with monosomal karyotype at the far end of the unfavorable prognostic spectrum. Haematologica-the Hematology Journal 96:491-493, 2011
65. Foran JM: New prognostic markers in acute myeloid leukemia: perspective from the clinic. Hematology Am Soc Hematol Educ Program 2010:47-55, 2010
66. Kiyoi H, Naoe T, Nakano Y, et al: Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. Blood 93:3074-80, 1999
67. Schnittger S, Schoch C, Dugas M, et al: Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. Blood 100:59-66, 2002
68. Thiede C, Steudel C, Mohr B, et al: Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. Blood 99:4326-35, 2002
69. Mead AJ, Gale RE, Hills RK, et al: Conflicting data on the prognostic significance of FLT3/TKD mutations in acute myeloid leukemia might be related to the incidence of biallelic disease. Blood 112:444-5; author reply 445, 2008
70. Daver N, Cortes J: Molecular targeted therapy in acute myeloid leukemia. Hematology 17 Suppl 1:S59-62, 2012
71. Metzelder S, Wang Y, Wollmer E, et al: Compassionate use of sorafenib in FLT3-ITD-positive acute myeloid leukemia: sustained regression before and after allogeneic stem cell transplantation. Blood 113:6567-71, 2009
72. Falini B, Mecucci C, Tiacci E, et al: Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med 352:254-66, 2005
73. Redner RL, Rush EA, Faas S, et al: The t(5;17) variant of acute promyelocytic leukemia expresses a nucleophosmin-retinoic acid receptor fusion. Blood 87:882-6, 1996
74. Schlenk RF, Dohner K, Krauter J, et al: Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. N Engl J Med 358:1909-18, 2008
75. Marcucci G, Mrozek K, Ruppert AS, et al: Prognostic factors and outcome of core binding factor acute myeloid leukemia patients with t(8;21) differ from those of patients with inv(16): a Cancer and Leukemia Group B study. Journal of Clinical Oncology 23:5705-17, 2005
76. Vardiman JW, Thiele J, Arber DA, et al: The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 114:937-51, 2009
77. Neubauer A, Maharry K, Mrozek K, et al: Patients with acute myeloid leukemia and RAS mutations benefit most from postremission high-dose cytarabine: A Cancer and Leukemia Group B study. Journal of Clinical Oncology 26:4603-4609, 2008

78. Gilliland DG, Griffin JD: The roles of FLT3 in hematopoiesis and leukemia. Blood 100:1532-1542, 2002
79. Murati A, Brecqueville M, Devillier R, et al: Myeloid malignancies: mutations, models and management. Bmc Cancer 12, 2012
80. Papaemmanuil E, Cazzola M, Boultwood J, et al: Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 365:1384-95, 2011
81. Plass C, Yu F, Yu L, et al: Restriction landmark genome scanning for aberrant methylation in primary refractory and relapsed acute myeloid leukemia; involvement of the WIT-1 gene. Oncogene 18:3159-65, 1999
82. Cameron EE, Baylin SB, Herman JG: p15(INK4B) CpG island methylation in primary acute leukemia is heterogeneous and suggests density as a critical factor for transcriptional silencing. Blood 94:2445-51, 1999
83. Wong IH, Ng MH, Huang DP, et al: Aberrant p15 promoter methylation in adult and childhood acute leukemias of nearly all morphologic subtypes: potential prognostic implications. Blood 95:1942-9, 2000
84. Mizuno S, Chijiwa T, Okamura T, et al: Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. Blood 97:1172-9, 2001
85. Ley TJ, Ding L, Walter MJ, et al: DNMT3A mutations in acute myeloid leukemia. N Engl J Med 363:2424-33, 2010
86. Yan XJ, Xu J, Gu ZH, et al: Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nature Genetics 43:309-U51, 2011
87. Ribeiro AFT, Pratcorona M, Erpelinck-Verschueren C, et al: Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. Blood 119:5824-5831, 2012
88. Ooi SKT, Qiu C, Bernstein E, et al: DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714-U13, 2007
89. Shih AH, Abdel-Wahab O, Patel JP, et al: The role of mutations in epigenetic regulators in myeloid malignancies. Nat Rev Cancer 12:599-612, 2012
90. Tadokoro Y, Ema H, Okano M, et al: De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. Journal of Experimental Medicine 204:715-22, 2007
91. Challen GA, Sun DQ, Jeong M, et al: Dnmt3a is essential for hematopoietic stem cell differentiation. Nature Genetics 44:23-U43, 2012
92. Hayette S, Thomas X, Jallades L, et al: High DNA Methyltransferase DNMT3B Levels: A Poor Prognostic Marker in Acute Myeloid Leukemia. Plos One 7, 2012
93. Monteferrario D, Noordermeer SM, Bergevoet S, et al: High DNA-methyltransferase 3B expression predicts poor outcome in acute myeloid leukemia, especially among patients with co-occurring NPM1 and FLT3 mutations. Blood Cancer J 4:e233, 2014
94. Dang CV: Links between metabolism and cancer. Genes Dev 26:877-90, 2012
95. Yan H, Parsons DW, Jin G, et al: IDH1 and IDH2 mutations in gliomas. N Engl J Med 360:765-73, 2009
96. Mardis ER, Ding L, Dooling DJ, et al: Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med 361:1058-66, 2009
97. Ward PS, Patel J, Wise DR, et al: The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 17:225-34, 2010
98. Dang L, White DW, Gross S, et al: Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462:739-44, 2009
99. Gross S, Cairns RA, Minden MD, et al: Cancer-associated metabolite

2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. Journal of Experimental Medicine 207:339-44, 2010

100. Sellner L, Capper D, Meyer J, et al: Increased levels of 2-hydroxyglutarate in AML patients with IDH1-R132H and IDH2-R140Q mutations. European Journal of Haematology 85:457-9, 2010
101. Moroni I, Bugiani M, D'Incerti L, et al: L-2-hydroxyglutaric aciduria and brain malignant tumors: a predisposing condition? Neurology 62:1882-4, 2004
102. Struys EA: D-2-Hydroxyglutaric aciduria: Unravelling the biochemical pathway and the genetic defect. Journal of Inherited Metabolic Disease 29:21-29, 2006
103. Struys EA, Salomons GS, Achouri Y, et al: Mutations in the D-2-hydroxyglutarate dehydrogenase gene cause D-2-hydroxyglutaric aciduria. American Journal of Human Genetics 76:358-360, 2005
104. Kranendijk M, Struys EA, van Schaftingen E, et al: IDH2 Mutations in Patients with D-2-Hydroxyglutaric Aciduria. Science 330:336-336, 2010
105. Figueroa ME, Abdel-Wahab O, Lu C, et al: Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18:553-67, 2010
106. Sasaki M, Knobbe CB, Munger JC, et al: IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. Nature 488:656-9, 2012
107. Iyer LM, Tahiliani M, Rao A, et al: Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle 8:1698-710, 2009
108. Loenarz C, Schofield CJ: Expanding chemical biology of 2-oxoglutarate oxygenases. Nature Chemical Biology 4:152-6, 2008
109. Tsukada Y, Fang J, Erdjument-Bromage H, et al: Histone demethylation by a family of JmjC domain-containing proteins. Nature 439:811-6, 2006
110. Schneider J, Shilatifard A: Histone demethylation by hydroxylation: chemistry in action. Acs Chemical Biology 1:75-81, 2006
111. Trewick SC, McLaughlin PJ, Allshire RC: Methylation: lost in hydroxylation? Embo Reports 6:315-320, 2005
112. Xu W, Yang H, Liu Y, et al: Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of alpha-Ketoglutarate-Dependent Dioxygenases. Cancer Cell 19:17-30, 2011
113. Lu C, Ward PS, Kapoor GS, et al: IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature 483:474-U130, 2012
114. Lane SW, Scadden DT, Gilliland DG: The leukemic stem cell niche: current concepts and therapeutic opportunities. Blood 114:1150-1157, 2009
115. Bernt KM, Armstrong SA: Targeting epigenetic programs in MLL-rearranged leukemias. Hematology Am Soc Hematol Educ Program 2011:354-60, 2011
116. Tkachuk DC, Kohler S, Cleary ML: Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell 71:691-700, 1992
117. Gu Y, Nakamura T, Alder H, et al: The T(4-11) Chromosome-Translocation of Human Acute Leukemias Fuses the All-1 Gene, Related to Drosophila-Trithorax, to the Af-4 Gene. Cell 71:701-708, 1992
118. Milne TA, Briggs SD, Brock HW, et al: MLL targets SET domain methyltransferase activity to Hox gene promoters. Molecular Cell 10:1107-1117, 2002
119. Nakamura T, Mori T, Tada S, et al: ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Molecular Cell 10:1119-1128, 2002
120. Krivtsov AV, Armstrong SA: MLL translocations, histone modifications

and leukaemia stem-cell development. Nature Reviews Cancer 7:823-833, 2007

121. Xu F, Li X, Wu LY, et al: Overexpression of the EZH2, RING1 and BMI1 genes is common in myelodysplastic syndromes: relation to adverse epigenetic alteration and poor prognostic scoring. Annals of Hematology 90:643-653, 2011

122. Morin RD, Johnson NA, Severson TM, et al: Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nature Genetics 42:181-U124, 2010

123. Sneeringer CJ, Scott MP, Kuntz KW, et al: Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. Proc Natl Acad Sci U S A 107:20980-5, 2010

124. Nikoloski G, Langemeijer SM, Kuiper RP, et al: Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nature Genetics 42:665-7, 2010

125. Abdel-Wahab O, Pardanani A, Patel J, et al: Concomitant analysis of EZH2 and ASXL1 mutations in myelofibrosis, chronic myelomonocytic leukemia and blast-phase myeloproliferative neoplasms. Leukemia 25:1200-2, 2011

126. Ernst T, Chase AJ, Score J, et al: Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nature Genetics 42:722-6, 2010

127. Makishima H, Jankowska AM, Tiu RV, et al: Novel homo- and hemizygous mutations in EZH2 in myeloid malignancies. Leukemia 24:1799-804, 2010

128. Dolnik A, Engelmann JC, Scharfenberger-Schmeer M, et al: Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. Blood 120:e83-92, 2012

129. Dawson MA, Kouzarides T: Cancer epigenetics: from mechanism to therapy. Cell 150:12-27, 2012

130. Welch JS, Ley TJ, Link DC, et al: The origin and evolution of mutations in acute myeloid leukemia. Cell 150:264-78, 2012

131. Papaemmanuil E, Gerstung M, Malcovati L, et al: Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 122:3616-27; quiz 3699, 2013

132. Kaminskas E, Farrell AT, Wang YC, et al: FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. Oncologist 10:176-82, 2005

133. Wilson VL, Jones PA, Momparler RL: Inhibition of DNA methylation in L1210 leukemic cells by 5-aza-2′-deoxycytidine as a possible mechanism of chemotherapeutic action. Cancer Research 43:3493-6, 1983

134. Fathi AT, Abdel-Wahab O: Mutations in epigenetic modifiers in myeloid malignancies and the prospect of novel epigenetic-targeted therapy. Adv Hematol 2012:469592, 2012

135. Itzykson R, Kosmider O, Cluzeau T, et al: Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. Leukemia 25:1147-52, 2011

136. Pollyea DA, Raval A, Kusler B, et al: Impact of TET2 mutations on mRNA expression and clinical outcomes in MDS patients treated with DNA methyltransferase inhibitors. Hematol Oncol 29:157-60, 2011

137. Sekeres MA, Tiu RV, Komrokji R, et al: Phase 2 study of the lenalidomide and azacitidine combination in patients with higher-risk myelodysplastic syndromes. Blood 120:4945-51, 2012

138. McCabe MT, Graves AP, Ganji G, et al: Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc Natl Acad Sci U S A 109:2989-94, 2012

139. Qi W, Chan H, Teng L, et al: Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. Proc Natl Acad Sci U S A 109:21360-5, 2012

140. Knutson SK, Wigle TJ, Warholic NM, et al: A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. Nature Chemical Biology 8:890-6, 2012

Chapter 2

Clinical and biological impact of *TET2* mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial

Mariam G. Aslanyan, Leonie I. Kroeze, Saskia M.C. Langemeijer, Theresia N. Koorenhof-Scheele, Marion Massop, Patricia van Hoogen, Ellen Stevens-Linders, Louis T. van de Locht, Evelyn Tönnissen, Adrian van der Heijden, Pedro da Silva-Coelho, Daniela Cilloni, Giuseppe Saglio, Jean-Pierre Marie, Ruoping Tang, Boris Labar, Sergio Amadori, Petra Muus, Roel Willemze, Erik W.A. Marijt, Theo de Witte, Bert A. van der Reijden, Stefan Suciu, Joop H. Jansen

This chapter has been adapted from:
Annals of Hematology, 2014 Aug;93(8):1401-12.

ABSTRACT

We assessed the prognostic impact of *TET2* mutations and mRNA expression in a prospective cohort of 357 adult AML patients <60 years of age enrolled in the EORTC/GIMEMA AML-12 06991 clinical trial. In addition the co-occurrence with other genetic defects and the functional consequences of *TET2* mutations were investigated. *TET2* mutations occurred in 7.6% of the patients and were an independent marker of poor prognosis (p=0.024). *TET2* and *IDH1/2* mutations strongly associated with aberrations in the DNA methyltransferase *DNMT3A*. Functional studies confirmed previous work that neither nonsense truncations, nor missense TET2 mutations induced 5-hydroxymethylcytosine formation. In addition, we now show that mutant *TET2* forms did not act in a dominant-negative manner when co-expressed with the wild-type protein. Finally, as loss-of-function *TET2* mutations predicted poor outcome, we questioned whether low *TET2* mRNA expression in cases of AML without *TET2* mutations would affect overall survival. Notably, also AML patients with low *TET2* mRNA expression levels showed inferior overall survival.

INTRODUCTION

Recently, a number of novel recurrent alterations have been described in adult AML, among which a group of epigenetic modifiers including *TET2*, *ASXL1*, *BCOR*, *DNMT3A*, *IDH1* and *IDH2* [1,2]. Verifying the relevance of these new genes as prognostic markers is crucial before they can be implemented into clinical practice to improve risk-adapted treatment. Understanding the molecular biological consequences of genetic mutations allows the development and application of targeted forms of therapy, aiming at the impaired proteins or the biological pathways they are involved in. *TET2* aberrations occur in various myeloid neoplasms. Mutations were found in 20-25% of myelodysplastic syndromes [3,4], 10-15% of myeloproliferative neoplasms [5], and up to 50% of chronic myelomonocytic leukemias [6]. In AML, *TET2* mutations affect 7-10% of the adulthood cases [3-5] and 1.5-4% of pediatric cases [7,8]. The impact of TET2 mutations on clinical outcome in AML is still controversial. In some studies, *TET2* mutations correlated with inferior survival, particularly in the subgroup of cytogenetically normal AML, whereas in other studies no prognostic impact of *TET2* mutations was observed [9-15]. So far, few prospectively collected cohorts have been analyzed for the impact of *TET2* mutations using multivariate analysis, and including the currently established genetic and other prognostic parameters [10-12, 14].

Several studies based on mouse models have shown that inactivation of *TET2* in myeloid cells induces a myeloproliferative syndrome, confirming the role of *TET2* mutations in the pathogenesis of myeloid malignancies [16-18]. The TET protein family comprising TET1, TET2 and TET3, belongs to the group of 2 oxoglutarate (2OG)-and Fe(II)-dependent dioxygenases, and is involved in a novel epigenetic modification, namely the conversion of 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC) [19]. TET proteins further oxidize 5hmC to formyl- and carboxylcytosine (5fC and 5caC). 5fC and 5caC are recognized by the thymine-DNA glycosylase (TDG), and replaced by unmodified cytosines through the DNA repair machinery [20,21]. Mapping of TET protein binding sites and 5hmC on the genome has shown that they localize to the transcription start sites of genes, suggesting a role for the TETs in the regulation of gene expression through modification of chromatin at promoter regions [22-24].

In this study, we assessed the incidence and clinical impact of *TET2* mutations in a prospective cohort of 357 AML patients who were treated uniformly according to the EORTC/GIMEMA AML-12 06991 clinical trial protocol. We determined the co-occurrence of *TET2* mutations with other molecular markers. In addition, the function of different *TET2* isoforms and TET2 mutants was tested in cell line models. Finally, mRNA levels of *TET2* were measured and correlated to survival.

MATERIALS AND METHODS

Patients

Bone marrow samples from 357 patients included in the EORTC/GIMEMA 06991 AML-12 [25] clinical trial NCT0004128 (1999- 2009) were obtained at diagnosis. Informed consent was obtained from all patients included in the study. To avoid bias, consecutive patients from the larger participating centers were included in the analysis. Median age was 46 years (range 15-60 years) and median follow-up was 5.63 years. The study was approved by the Institutional Review Board of the Radboud University Nijmegen Medical Centre and all aspects of this study abided by the principles laid out in the declaration of Helsinki. All patients had untreated newly diagnosed AML, ≥ 30% blast cells in bone marrow smears, and represented all AML FAB subtypes, except AML-M3. Briefly, the AML-12 trial assessed the value of high vs. standard dose Ara-C during induction, as well as IL-2 after intensive consolidation/auto-SCT.

Q-PCR and mutational analysis

For Q-PCR analysis RNA was isolated from healthy donor-derived G-SCF-mobilized CD34+ cells, or patient-derived bone marrow or peripheral blood mononuclear cells. mRNA levels of *TET2* isoform 1 and isoform 2, *TET1* and *TET3* were determined as previously described [3,4]. DNA sequencing of *TET2* isoform 1 (NM_001127208.2) was performed as described [3]. cDNA sequencing primers are listed in **Supplementary Table S1**. *EVI1*, *IDH1* R132, *IDH2* R140 and R172 detection was performed as described [26,27]. Direct DNA sequencing of *DNMT3A* (ENST00000264709) was performed on exons 7-23 either by Sanger sequencing (primer sequences listed in **Supplementary Table S2**) or 454-based next generation sequencing (454 Life Sciences, Branford, CT, USA) [28].NGS data were analyzed using R453Plus1Toolbox, GS Variant Analyzer Software 2.6 (454 Life Sciences, Branford, CT) and Sequence Pilot version 3.5.2 (JSI Medical Systems, Kippenheim, Germany)[29]. NPM1 and FLT3 mutation analyses were developed previously [30,31].

Statistical analysis

Overall survival (OS) was computed from the date of registration in the EORTC 06991 study until death (whatever the cause) or until last follow-up (censored observation). Survival distribution of OS was estimated according to the Kaplan-Meier technique. The Cox Proportional Hazards Model was used to determine the prognostic importance of each factor analyzed and to obtain hazard ratio (HR) estimates as well as corresponding 95% confidence intervals (95% CI). All statistical analyses were performed according to the intention-to-treat principle. Being a prognostic factor analysis, only eligible patients were included. For contingency tables, the significance of a relationship between categorical 2 variables, the Fisher exact test was used. Statistical analyses were performed by using SAS® (Cary, USA).

Plasmids used for overexpression in cell line models

N-terminally tagged eGFP-TET2 isoform 1 plasmid was kindly provided by Dr. O. Bernard. *TET2* isoform 1 (NM_001127208.2), *TET2* isoform 2 (NM_017628.4), TET2 catalytic domain (a.a. 1104 – a.a. 2002), TET2 box 2 (a.a. 1789-a.a. 2002) sequences were cloned into pDONR201 using the Gateway cloning technology (Life technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. The entry vectors were further transferred into a Gateway based pDest733 [32] {Relaix, 2004 #374}destination vector to create N-terminal mRFP fusion proteins.

TET2 mutant sequences were generated using the QuikChange XL site-directed mutagenesis kit (Stratagene, Agilent technologies, CA, USA) as recommended by the manufacturers and cloned into pDONR201. The entry vectors were used for transfer into pDest733 and pDest501 Gateway-based expression plasmids [35] to create N-terminal mRFP and eCFP fusions, respectively.

Cell culture, transfection and immunofluorescence

COS-1 cells were grown in Iscove's Modified Dulbecco's Media (IMDM) (Life technologies) supplemented with 10% fetal calf serum and penicillin and streptomycin. Cells were transfected with Lipofectamin 2000 reagent (Life technologies). Proteins were expressed for 20-24h. Subsequently, cells were fixed with 4% PFA for 20 min, washed extensively with PBS and permeabilized with 0.5% Triton X-100 in PBS for 5 min. To allow detection of 5hmC, cells were treated with 2M HCl for 20-25 min. HCl was neutralized with 100mM Tris-HCl for 10 min (pH=8), followed by extensive washing with PBS. A blocking step was performed with 5%BSA in PBS/Tween for 30 min. Afterwards, samples were incubated with a primary antibody against 5hmC (Active motif, CA, USA; 1:500 dilution) for 2h at RT, and subsequently washed with PBS. To allow primary antibody detection, the following secondary antibodies were used: Alexa 488 (Life technologies; 1:500 dilution) or Alexa 350 (Life technologies; 1:50 dilution). Samples were mounted with Vectashield with DAPI or Antifade Prolong Gold (in the case of Alexa 350). Samples were analyzed with a Leica DMBRE (type 301-371.011) microscope. Images were taken at 40x or 100x magnification using a DC350F (10447116) camera and Leica FW4000 acquisition software (Leica Microsystems, Wetzlar, Germany).

RESULTS

TET2 mutations correlate with poor overall survival in AML in univariate analysis

Bidirectional Sanger sequencing of the entire *TET2* coding region, and the splice donor and acceptor sites, revealed mutations in 27/357 (7.6%) of the patients (for detailed patient characteristics see **Table 1;** localization of mutations is depicted in **Supplementary Fig. S1**). In 19 of the cases, mutations affected one allele, whereas

Table 1. Clinical and molecular characteristics according to TET2 mutation status in 357 adult AML patients

		TET2 mutation		Correlation to *TET2* mutation	Hazard ratio	p-value
		No (N=330)	Yes (N=27)	(p-value)	(95% CI)	(Wald test)
		N (%)	N (%)			
Type of AML	de novo	311 (94.2)	27 (100)	0.38		
	secondary	19 (5.8)	0 (0.0)			
Sex	Male	160 (48.5)	13 (48.1)	0.97		
	Female	170 (51.5)	14 (51.9)			
Age (years) at diagnosis	Median	45.0	54.0			
	Range	15.0 - 60.0	24.0 - 59.0			
	15-25	36 (10.9)	1 (3.7)	0.015	1.0	0.05
	26-45	133 (40.3)	6 (22.2)		1.33 (0.79, 2.25)	
	46-60	161 (48.8)	20 (74.1)		1.71 (1.03, 2.85)	
WHO-ECOG performance status	PS 0	181 (54.8)	13 (48.1)	0.65	1.0	0.11
	PS 1	119 (36.1)	12 (44.4)		0.89 (0.66, 1.19)	
	PS 2	22 (6.7)	1 (3.7)		1.30 (0.76, 2.23)	
	PS 3-4	8 (2.4)	1 (3.7)		2.03 (0.99, 4.17)	
WBC (x10^9/l) at diagnosis	Median	21.5	32.1			
	Range	0.5 - 319.6	1.1 - 240.8			
	< 25	183 (55.5)	13 (48.1)	0.69	1.0	0.44
	25-99.9	105 (31.8)	10 (37.0)		1.16 (0.86, 1.56)	
	>= 100	42 (12.7)	4 (14.8)		1.09 (0.72, 1.64)	
FAB subtype	M0	13 (3.9)	1 (3.7)	0.06	1.0	0.16
	M1	87 (26.4)	2 (7.4)		1.21 (0.55, 2.65)	
	M2	93 (28.2)	9 (33.3)		1.18 (0.54, 2.59)	
	M4	60 (18.2)	3 (11.1)		0.91 (0.40, 2.08)	
	M5	65 (19.7)	11 (40.7)		1.53 (0.69, 3.37)	
	M6	6 (1.8)	1 (3.7)		2.23 (0.75, 6.64)	
	M7	1 (0.3)	0 (0.0)		4.72 (0.58, 38.64)	
	Unknown	5 (1.5)	0 (0.0)	-		
Cytogenetic risk group	Good	41 (12.4)	0 (0.0)	0.004	1.0	<0.0001
	Intermediate (NN,-Y)	111 (33.6)	15 (55.6)		2.63 (1.40, 4.97)	
	Others	73 (22.1)	1 (3.7)		3.17 (1.65, 6.11)	

Table 1. Clinical and molecular characteristics according to TET2 mutation status in 357 adult AML patients (*Continued*)

		TET2 mutation		Correlation to *TET2* mutation	Hazard ratio	p-value
		No (N=330)	Yes (N=27)	(p-value)	(95% CI)	(Wald test)
		N (%)	N (%)			
	Poor (-5/5q-,-7/7q-, Complex)	29 (8.8)	1 (3.7)		6.16 (3.01, 12.60)	
	UNK/ND/Failure	76 (23.0)	10 (37.0)	-		
*NPM1*mutation	No	224 (67.9)	13 (48.1)	0.03	1.0	0.67
	Yes	102 (30.9)	14 (51.9)		0.94 (0.70, 1.25)	
	Unknown	4 (1.2)	0 (0.0)	-		
FLT3-ITD	No	254 (77.0)	19 (70.4)	0.48	1.0	<0.0001
	Yes	73 (22.1)	8 (29.6)		2.39 (1.78, 3.20)	
	Unknown	3 (0.9)	0 (0.0)	-		
FLT3-TKD	No	294 (89.1)	22 (81.5)	0.31	1.0	0.19
	Yes	30 (9.1)	4 (14.8)		0.73 (0.46, 1.17)	
	Unknown	6 (1.8)	1 (3.7)	-		
IDH1 mutation	No	288 (87.3)	24 (88.9)	0.24	1.0	0.08
	Yes	28 (8.5)	0 (0.0)		1.50 (0.96, 2.36)	
	Unknown	14 (4.2)	3 (11.1)	-		
IDH2 mutation	No	278 (84.2)	22 (81.5)	1.0	1.0	0.49
	Yes	36 (10.9)	2 (7.4)		0.86 (0.55, 1.33)	
	Unknown	16 (4.8)	3 (11.1)	-		
IDH mutation	No	254 (77.0)	22 (81.5)	0.28	1.0	
	Yes	62 (18.8)	2 (7.4)		1.07 (0.76, 1.51)	0.69
	Unknown	14 (4.2)	3 (11.1)	-		
EVI1 overexpression	No	103 (31.2)	7 (25.9)	1.0	1.0	0.0004
	Yes	7 (2.1)	0 (0.0)		3.94 (1.75, 8.88)	
	Unknown	220 (66.7)	20 (74.1)	-		
DNMT3A mutation	No	50 (15.2)	6 (22.2)	0.02	1.0	0.04
	Yes	18 (5.5)	9 (33.3)		1.73 (1.01, 2.99)	
	Unknown	262 (79.4)	12 (44.4)	-		

Abbreviations: CI, confidence interval; WBC, white blood cell count; PS, performance status; ITD, internal tandem duplication; TKD, tyrosine kinase domain

in 8 patients two *TET2* aberrations co-existed, suggesting bi-allelic inactivation (for detailed patient characteristics of *TET2* mutated patients see **Table 2**). In univariate analysis, the various cytogenetic subtypes showed the expected correlations with overall survival, with the t(8;21) and inv(16) translocations correlating with a good prognosis, and -5, -7 and complex karyotypes signifying poor outcome (**Supplementary Fig. S2A**). *NPM1* mutations did however not have prognostic impact (**Supplementary Fig. S2B**). In line with published data, the presence of a *FLT3*-ITD mutation correlated with inferior clinical outcome (p<0.0001), whereas a *FLT3*-TKD mutation associated with intermediate or slightly better survival (**Supplementary Fig. S2C and S2D**). Furthermore, *EVI1* overexpression was strongly associated with inferior prognosis (**Supplementary Fig. S2E**, p=0.0004).

The presence of a *TET2* mutation correlated with poor overall survival (**Fig. 1A**, p=0.025). *IDH2* mutations did not show a significant effect on overall survival (**Fig. 1C**), whereas a trend towards poor prognosis was present in patients with an *IDH1* mutation (**Fig. 1D**, p=0.075). Finally, *DNMT3A* mutations also predicted unfavorable outcome (**Fig. 1B**, p=0.04).

Table 2. Characteristics of TET2 mutated patients. Additional molecular markers that were assessed are: FLT3-ITD, FLT3-TKD, IDH1, IDH2, DNMT3A, NPM1 mutation, and EVI-1 overexpression.

UPN[1]	*TET2* mutation by sequencing			Cytogenetics	*FLT3*-ITD, *FLT3*-TKD, *IDH1*, *IDH2*, *DNMT3A*, *NPM1*, *EVI1* overexpression
	Nonsense	Frameshift/ In-frame deletion	Missense/ Splice site		
36			p.S1039L	46,XX	IDH2 R172L, DNMT3A p.F640Y
98			p.R1359H	46,XY	NPM1
171			p.A31T	46,XX	-
329		p.P1115LfsX2		46,XX	NPM1, FLT3-ITD, DNMT3A IVS9-1G>T
356		p.N275IfsX18		46,XX	FLT3-ITD, DNMT3A p.R882C
362[2]			p.R1896M	46,XX,t(4;7)(q24;q21) [19]/ 46,XX [1]	NPM1, FLT3-TKD
379	p.Q255X			Failure	-
404	p.Q414X			Unknown	-
729			p.S358R	46,XY	NPM1
791			p.I1873N	45,XY,-7,del(20)(q11)	-
806[2]		p.P929LfsX24 p.S137VfsX8		46,XY	NPM1, FLT3-ITD, DNMT3A p.D252X
858			p.L956V	Failure	NPM1, FLT3-ITD
879			p.Q742K	Unknown	NPM1, FLT3-ITD

Table 2. Characteristics of TET2 mutated patients. Additional molecular markers that were assessed are: FLT3-ITD, FLT3-TKD, IDH1, IDH2, DNMT3A, NPM1 mutation, and EVI-1 overexpression. (*Continued*)

UPN[1]	*TET2* mutation by sequencing			Cytogenetics	*FLT3*-ITD, *FLT3*-TKD, *IDH1*, *IDH2*, *DNMT3A*, *NPM1*, *EVI1* overexpression
	Nonsense	Frameshift/ In-frame deletion	Missense/ Splice site		
893		p.I750_K753		46,XY	-
1062			p.I1873N	46,XY	-
1266[2]			p.G1869E IVS2-2A>G	46,XX	NPM1, FLT3-TKD, DNMT3A (p.N501S; p.R882H)
1398			p.H850R	46,XY	NPM1, FLT3-ITD, DNMT3A p.R882H
1456		p.Q831TfsX15		46,XX	FLT3-ITD, DNMT3A p.R882H
1552[2]	p.Q917X			Failure	-
1612[2]			p.C1378R p.P1419R	Failure	-
1744			p.R1572W	Failure	-
1832		p.T621SfsX13		46,XX	NPM1, FLT3-TKD
1879			p.T1397I	46,XX	NPM1
1915[2]	p.S1758X	p.T1554SfsX16		Failure	NPM1, FLT3-TKD
1928[2]	p.Q749X		p.D1384N	46,XY	NPM1, DNMT3A IVS9del -18_+1
1965			p.S1246L	Failure	NPM1, FLT3-ITD, DNMT3A p.R882H, IDH2 R140Q
2112[2]			p.R1261H p.R1359S	Failure -> FISH: MLL (11q23) rearrangement negative; 1 extra MLL signal [44/150] and 2 extra MLL signals [74/150]	-

1 UPN indicates unique patient number
2 indicates patients with more than one affected TET2 allele, based on the occurrence of two separate TET2 mutations, or a homozygous mutation as observed in the Sanger sequencing results, pointing to LOH by (micro) deletions or uniparental disomy.

Association of TET2 mutations with other prognostic features including molecular and cytogenetic markers

TET2 mutations correlated with older age, and were almost exclusively present in the group of cytogenetically normal AML (Table 1). We found no association between TET2 mutations and performance status or WBC at diagnosis. There was, however, an association between TET2 mutations and FAB subtype (p=0.06). TET2 mutations were less frequent in AML-M1, and enriched in AML-M5 patients. NPM1 and TET2 defects were significantly correlated (p=0.03), whereas FLT3-ITD

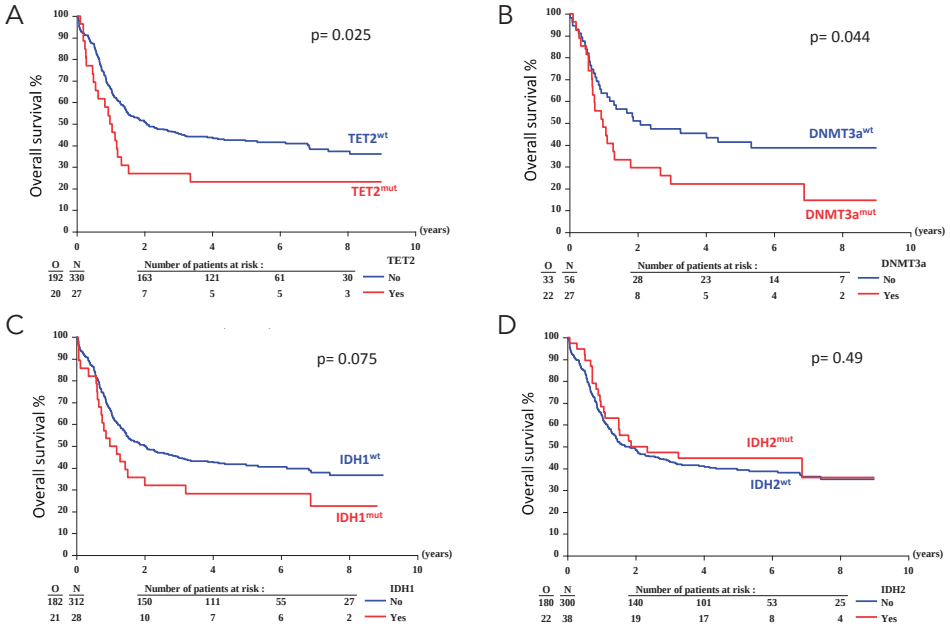


Figure 1. *Univariate analysis assessing the effect of mutations in genes implicated in the DNA (hydroxy)methylation pathway. TET2 (A) DNMT3A (B) IDH1 (C) and IDH2 (D).*

and FLT3-TKD aberrations were often present together with TET2 mutations but their co-occurrence did not reach statistical significance. TET2 and IDH mutations seldom co-existed in the same patient as may be expected since IDH mutations abrogate the activity of TET2 [33]. Notably, both TET2 and IDH mutations strongly correlated with the presence of a DNMT3A mutation (p=0.02 and p=0.004 respectively). Patients who carried DNMT3A aberrations together with either a TET2 or an IDH mutation, very frequently had an additional FLT3-ITD.

TET2 mutations represent an independent marker of poor prognosis in AML

Cox regression analysis was performed to determine whether TET2 mutations represent an independent marker of poor prognosis. In multivariate analysis, adjusting for age, cytogenetic/FLT3-ITD risk group, NPM1 mutation, AML type, institution, and stratifying by treatment arm, TET2 mutations remained an independent predictor of inferior survival (p=0.01, HR = 1.86, 95% CI: 1.15 – 3.00. Within the poor-risk group of patients carrying a FLT3-ITD and/or a high-risk cytogenetic profile (-5, -7, complex), which conferred the expected poor overall survival, the presence of an additional TET2 mutation defined a subgroup of patients with an even more dismal prognosis (**Fig. 2A,** p=0.033). NPM1 mutations had only a minor effect towards favorable outcome in the cytogenetically normal

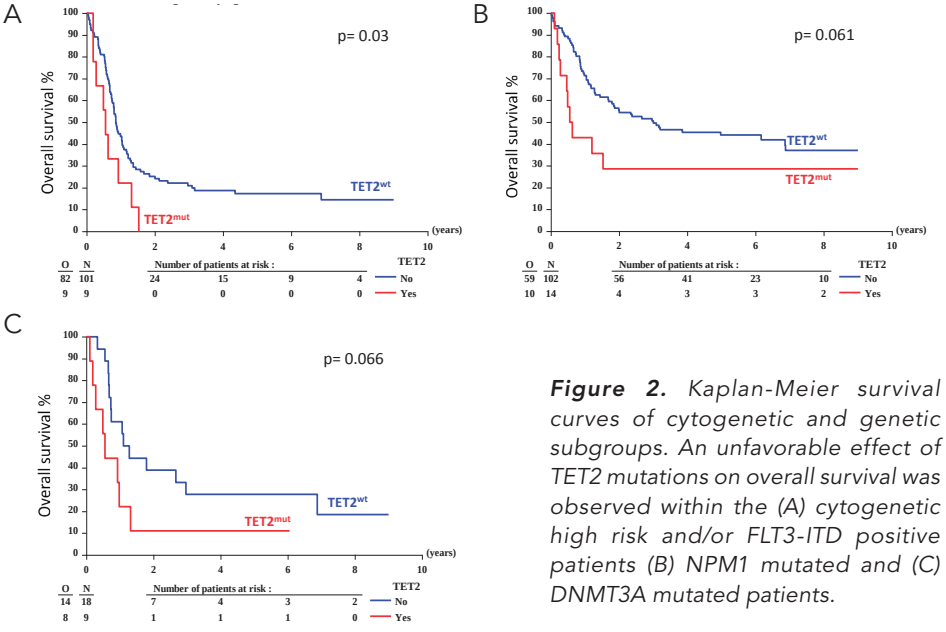


Figure 2. *Kaplan-Meier survival curves of cytogenetic and genetic subgroups. An unfavorable effect of TET2 mutations on overall survival was observed within the (A) cytogenetic high risk and/or FLT3-ITD positive patients (B) NPM1 mutated and (C) DNMT3A mutated patients.*

subgroup without FLT3-ITD (**Supplementary Fig. 2B**). Also within the NPM1 mutated group, TET2 mutations conferred a trend towards worse survival (**Fig. 2B**, p=0.061). Finally, as DNMT3A and TET2 mutations frequently co-occurred (**Table 1**), and DNMT3A mutations correlated with inferior survival (**Fig. 1B**), we tested what the effect of TET2 mutations would be in the context of a concomitant DNMT3A mutation. Also in the DNMT3A mutated subgroup, the presence of a TET2 mutation further exacerbated the poor prognosis of patients (**Fig. 2C**, p=0.066). Treatment with standard or high dose Ara-C did not alter the effect of TET2 mutation, as these mutations remained a poor indicator of prognosis in both patient groups (**Supplementary Fig. S3**).

Functionality of different TET2 splice- and truncation variants

Although studies so far have always focused on the longest isoform of TET2 (isoform 1), there are three major TET2 transcripts that may be expressed in cells, as we previously showed [3] (**Supplementary Fig. S4A**). The shorter isoforms lack the catalytic domain and their function remains unknown. The expression levels of isoform 1, and of the shorter form, isoform 2, are comparable. Furthermore, isoform 2 is reminiscent of some of the truncation mutants found in patients [3]. To investigate whether isoform 2 exhibits dominant-negative features, we tested its effect on the capacity of isoform 1 to generate 5hmC. Several expression constructs were created that contained TET2 isoform 1, TET2 isoform 2, the catalytic domain (CD), and the second conserved region of the catalytic domain (box 2) (**Supplementary**

Fig. S4B). To monitor expression, all proteins were N-terminally fused to mRFP (monomeric red fluorescent protein). COS-1 cells were transiently transfected with each of the expression vectors and immunostaining was performed using an antibody specific for 5-hydroxymethylcytosine. Expression of isoform 1 and the isolated catalytic domain resulted in a clear formation of 5hmC. Isoform 2 and the isolated, conserved box 2 showed no enzymatic activity (Supplementary Fig. S4C). Importantly, co-expression of isoform 2 with isoform 1 did not inhibit the generation of 5hmC (Supplementary Fig. S4D), implying that isoform 2 does not exhibit dominant-negative properties.

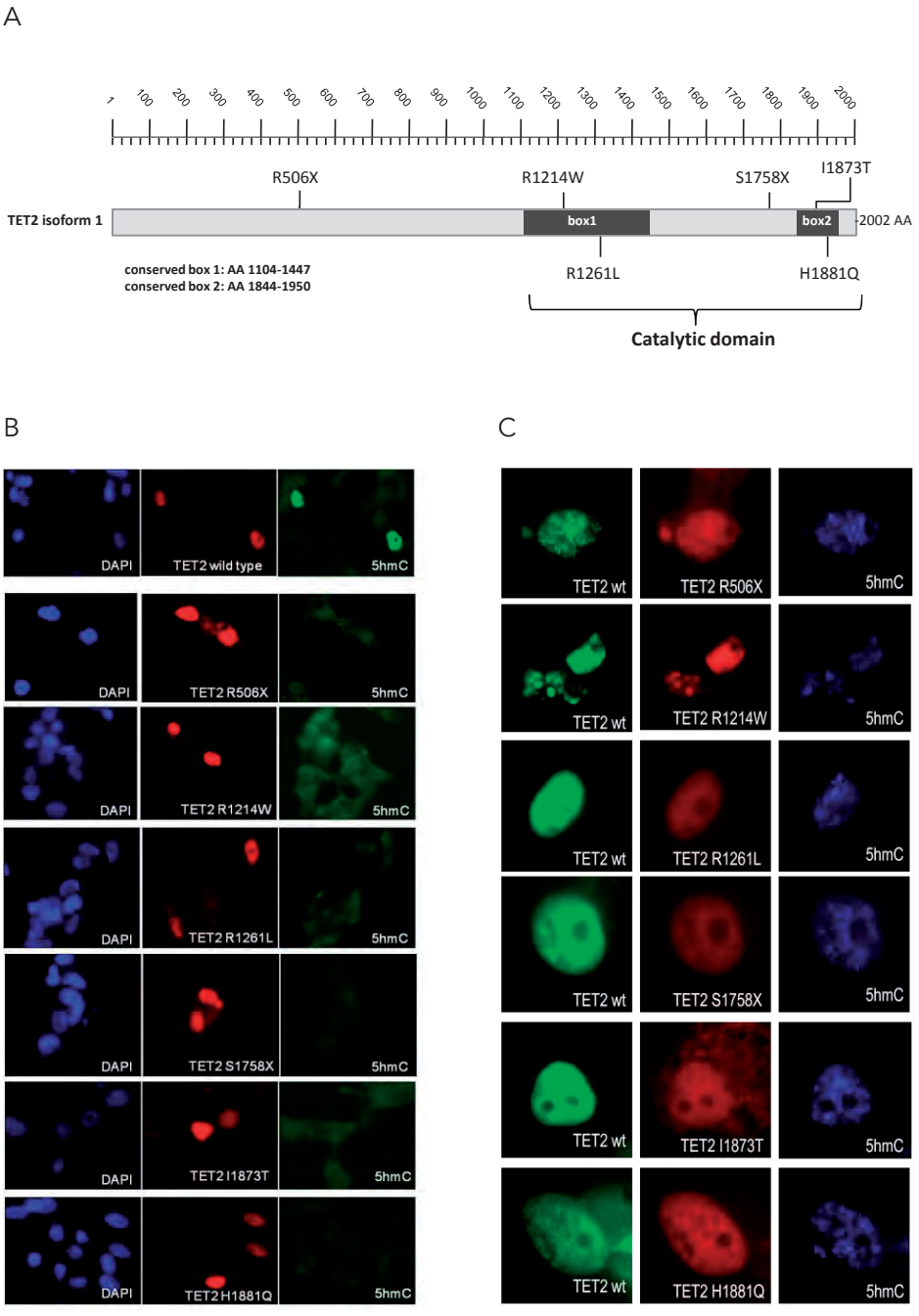
TET2 mutations result in loss of function and are not dominant-negative

Mutations in TET2 are scattered throughout the entire coding region of the gene and may result in the production of truncated proteins, as well as single amino acid substitutions. In-vitro studies and mouse models have suggested that likely all TET2 mutations lead to loss of function. Moreover, in patients, despite the presence of a wild type allele, DNA hydroxymethylation is strongly reduced [34]. It remains however possible that the mutant protein inhibits the function of the wild-type allele in a dominant-negative fashion. To test this, six mutant TET2 sequences were cloned into expression vectors, each N-terminally fused to mRFP. Two nonsense mutations were selected that delete either both conserved boxes (R506X)—hence the whole catalytic domain—or yield a protein containing only the first conserved box (box 1), causing a partial deletion of the catalytic domain (S1758X). Additionally, four missense substitutions were generated, two in box 1 (R1214W and R1261L) and two in box 2 (I1873T and H1881Q) (Fig. 3A). All six mutants, including the single amino acid substitutions, lost the capability to generate 5hmC (Fig. 3B). Furthermore, co-expression of GFP-tagged wild-type TET2 with each of the mutant proteins did not result in loss of hydroxymethylation, indicating that, with respect to the formation of 5hmC, the mutants are not dominant negative (Fig. 3C).

Effect of wild-type TET2 on survival of AML patients

As diminished TET2 function appears to be the mechanism through which mutations in this gene contribute to oncogenesis, and patients who carry TET2 mutations show a significantly worse overall survival, we tested whether the

Figure 3. *TET2 mutant proteins show impaired catalytic activity, but do not exhibit dominant-negative features when co-expressed with wild-type protein. (A) Localization of TET2 mutations found in patients with myeloid malignancies used for functional testing. Two of the mutants result in truncated proteins lacking either both conserved boxes within the catalytic region, or only the second conserved box, whereas the rest of the mutants result in missense substitutions in the conserved boxes. All aberrations leading to amino acid changes have been confirmed either by us or other as true acquired somatic mutations. The H1881Q mutation targets an Fe (II)-binding site in the catalytic domain of TET2. (B) Mutant forms of* ▶



TET2 found in patients are unable to produce 5hmC unlike wild-type protein. Constructs contain an N-terminal mRFP tag. Images were taken at 40x magnification. (C) mRFP-tagged mutant forms of TET2 do not act in a dominant-negative manner over wild-type GFP-TET2. Images were taken at 100x magnification

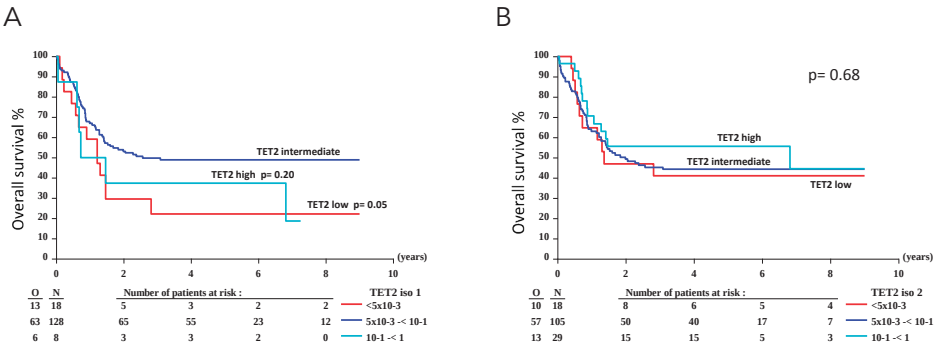


Figure 4. *Aberrant expression of the wild-type TET2 isoform 1 has a negative impact on overall survival (A), whereas the TET2 isoform 2 expression has no effect on survival (B).*

expression levels of isoforms 1 and 2 of TET2 may impact clinical outcome in patients without a TET2 or IDH mutation. The mRNA levels of both isoforms varied between different patients. No significant differences in expression were observed between patients with and without a TET2 mutation (**Supplementary Fig. S5A**). When compared to mRNA levels of TET2 in CD34+ cells, AML patients showed expression levels that were distributed in a wider range (**Supplementary Fig. S5B**). The median levels of TET2 expression in patients were similar to those measured in healthy CD34+ control cells. There was no difference in overall survival when TET2 isoform 2 expression was considered (**Fig. 4B**). In contrast, overall survival analysis showed that low TET2 isoform 1 expressers had an inferior clinical outcome (**Fig. 4A** p= 0.05).

DISCUSSION

The prognostic significance of TET2 mutations in AML patients has been controversial. Initially, Abdel-Wahab et al.[9] showed an association with poor OS in a smaller, retrospective study, whereas Nibourel et al. [13] reported no correlation between TET2 mutational status and clinical outcome, albeit in a pre-selected group of de novo AML patients achieving complete remission. Consequently, studies involving larger cohorts of patients were explored. Shen et al. [14] found no prognostic significance for TET2 mutations in 605 de novo AML patients who had no other recognizable karyotypic abnormalities except 11q23, and Metzeler et al. [12] identified TET2 mutations as a predictor of inferior survival in favorable-risk cytogenetically normal primary AML patients, as defined by the European LeukemiaNet (ELN) guidelines. Chou et al. [10] defined TET2 aberrations as a marker of poor prognosis in AML patients of intermediate risk, and Weismann et al. [15] also showed a negative impact of TET2 mutations on survival in favorable-risk

AML patients with normal cytogenetics. Most recently, [11] Gaidzik et al. did not find a prognostic impact of TET2 mutations in 330 cytogenetically normal AML patients. In the current study, we show a negative prognostic impact of TET2 mutations in a sizeable, prospectively studied cohort of adult AML patients <60 years of age. The negative effect of TET2 mutation on overall survival was present in various risk groups. In accordance with previous reports, TET2 mutations were significantly associated with older age. In addition, TET2 alterations were reported in normal elderly individuals with clonal hematopoiesis [35] and TET2 mutations were shown to be present in a pre-leukemic stage of leukemia evolution [36]. Furthermore, TET2 mutations are found in many different subtypes of myeloid malignancies. This suggests that cells carrying TET2 defects may clonally expand, but that additionally acquired mutations mostly determine the subtype and severity of the disease that eventually causes clinical problems.

It has previously been noted that TET2 and DNMT3A mutations co-occur in T-cell lymphomas [37]. In AML, a correlation between IDH and DNMT3A has been observed in AML [14]. In the present study, we show that both TET2 and IDH mutations were enriched in patients who also carried a DNMT3A defect. Although DNMT3A mutations conferred poor prognosis (in line with previous work [38]), the additional presence of TET2 alterations in these patients defined a subgroup with a particularly unfavorable outcome. As established earlier, TET2 and IDH mutations were mutually exclusive since both groups of mutations perturb a common biological pathway [33,39,40]. The high incidence of DNMT3A mutations in both of these groups is highly indicative of a cooperative mechanism through which mutations impairing DNA hydroxymethylation and DNA methylation contribute to leukemogenesis. Further functional characterization of this interaction in in vivo models is necessary and may provide important insights into targeted therapies aimed at the DNA (de)methylation pathway within patients with this particular genetic profile.

Testing of TET2 mutations in vitro [34], as well as in mouse models consistently points to a loss-of-function phenotype [16-18,41]. This is further supported in this manuscript. In addition, we show that TET2 mutants do not suppress the function of the wild-type protein, and hence do not show dominant-negative traits. As mRNA levels of TET2 in AML patients varied over a wide range compared to normal hematopoietic cells, this raised the possibility that the level of TET2 mRNA expression could be relevant in patients who do not carry loss-of-function mutations of the TET2 gene. Analysis of the effect of wild-type TET2 expression on clinical outcome in patients without TET2 or IDH mutations demonstrated that patients with low levels of TET2 expression had a worse overall survival. We conclude that both loss-of-function mutations and low expression of TET2 are markers in of poor prognosis in AML indicating that the development of therapies targeting the DNA hydroxymethylation pathway could be beneficial for these patients.

ACKNOWLEDGEMENTS

This work was supported by grants from the EORTC (EGAM 2009 NOCI), the EORTC Charitable Trust, and the Netherlands Institute for Regenerative Medicine (NIRM).

REFERENCES

1. Smith ML, Hills RK, Grimwade D (2011). Independent prognostic variables in acute myeloid leukaemia. Blood Rev 25(1): 39-51
2. Patel JP, Gönen M, Figueroa ME, Fernandez H, Sun Z, Racevskis J, Van Vlierberghe P, Dolgalev I, Thomas S, Aminova O, Huberman K, Cheng J, Viale A, Socci ND, Heguy A, Cherry A, Vance G, Higgins RR, Ketterling RP, Gallagher RE, Litzow M, van den Brink MR, Lazarus HM, Rowe JM, Luger S, Ferrando A, Paietta E, Tallman MS, Melnick A, Abdel-Wahab O, Levine RL. (2012) Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. N Engl J Med. 366(12):1079-1089.
3. Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, Stevens-Linders E, van Hoogen P, van Kessel AG, Raymakers RA, Kamping EJ, Verhoef GE, Verburgh E, Hagemeijer A, Vandenberghe P, de Witte T, van der Reijden BA, Jansen JH (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet 41(7): 838-842
4. Langemeijer SM, Aslanyan MG, Jansen JH (2009). TET proteins in malignant hematopoiesis. Cell Cycle 8(24): 4044-4048.
5. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Massé A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, Lécluse Y, Plo I, Dreyfus FJ, Marzac C, Casadevall N, Lacombe C, Romana SP, Dessen P, Soulier J, Viguié F, Fontenay M, Vainchenker W, Bernard OA (2009). Mutation in TET2 in myeloid cancers. N Engl J Med 360(22): 2289-2301
6. Kosmider O, Gelsi-Boyer V, Ciudad M, Racoeur C, Jooste V, Vey N, Quesnel B, Fenaux P, Bastie JN, Beyne-Rauzy O, Stamatoulas A, Dreyfus F, Ifrah N, de Botton S, Vainchenker W, Bernard OA, Birnbaum D, Fontenay M, Solary E. (2009). TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. Haematologica-the Hematology Journal 94(12): 1676-1681.
7. Langemeijer SMC, Jansen JH, Hooijer J, van Hoogen P, Stevens-Linders E, Massop M, Waanders E, van Reijmersdal SV, Stevens-Kroef MJ, Zwaan CM, van den Heuvel-Eibrink MM, Sonneveld E, Hoogerbrugge PM, van Kessel AG, Kuiper RP (2011). TET2 mutations in childhood leukemia. Leukemia 25(1): 189-192.
8. Liang DC, Liu HC, Yang CP, Jaing TH, Hung IJ, Yeh TC, Chen SH, Hou JY, Huang YJ, Shih YS, Huang YH, Lin TH, Shih LY (2013). Cooperating gene mutations in childhood acute myeloid leukemia with special reference on mutations of ASXL1, TET2, IDH1, IDH2 and DNMT3A. Blood 121(15): 2988-2995.
9. Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, Malinge S, Yao J, Kilpivaara O, Bhat R, Huberman K, Thomas S, Dolgalev I, Heguy A, Paietta E, Le Beau MM, Beran M, Tallman MS, Ebert BL, Kantarjian HM, Stone RM, Gilliland DG, Crispino JD, Levine RL (2009). Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood 114(1): 144-147.
10. Chou WC, Chou SC, Liu CY, Chen CY, Hou HA, Kuo YY, Lee MC, Ko BS, Tang JL, Yao M, Tsay W, Wu SJ, Huang SY, Hsu

SC, Chen YC, Chang YC, Kuo YY, Kuo KT, Lee FY, Liu MC, Liu CW, Tseng MH, Huang CF, Tien HF (2011). TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. Blood 118(14): 3803-3810.

11. Gaidzik VI, Paschka P, Späth D, Habdank M, Köhne CH, Germing U, von Lilienfeld-Toal M, Held G, Horst HA, Haase D, Bentz M, Götze K, Döhner H, Schlenk RF, Bullinger L, Döhner K (2012). TET2 Mutations in Acute Myeloid Leukemia (AML): Results From a Comprehensive Genetic and Clinical Analysis of the AML Study Group. Journal of Clinical Oncology 30(12): 1350-1357.
12. Metzeler KH, Maharry K, Radmacher MD, Mrózek K, Margeson D, Becker H, Curfman J, Holland KB, Schwind S, Whitman SP, Wu YZ, Blum W, Powell BL, Carter TH, Wetzler M, Moore JO, Kolitz JE, Baer MR, Carroll AJ, Larson RA, Caligiuri MA, Marcucci G, Bloomfield CD (2011). TET2 Mutations Improve the New European LeukemiaNet Risk Classification of Acute Myeloid Leukemia: A Cancer and Leukemia Group B Study. Journal of Clinical Oncology 29(10): 1373-1381.
13. Nibourel O, Kosmider O, Cheok M, Boissel N, Renneville A, Philippe N, Dombret H, Dreyfus F, Quesnel B, Geffroy S, Quentin S, Roche-Lestienne C, Cayuela JM, Roumier C, Fenaux P, Vainchenker W, Bernard OA, Soulier J, Fontenay M, Preudhomme C. (2010). Incidence and prognostic value of TET2 alterations in de novo acute myeloid leukemia achieving complete remission. Blood 116(7): 1132-1135.
14. Shen Y, Zhu YM, Fan X, Shi JY, Wang QR, Yan XJ, Gu ZH, Wang YY, Chen B, Jiang CL, Yan H, Chen FF, Chen HM, Chen Z, Jin J, Chen SJ (2011). Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. Blood 118(20): 5593-5603.
15. Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C, Dicker F, Fasan A, Haferlach C, Haferlach T, Kern W, Schnittger S, Kohlmann A. (2012) Landscape of TET2 mutations in acute myeloid leukemia. Leukemia. 26 (5) 934-942
16. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, Tsangaratou A, Rajewsky K, Koralov SB, Rao A (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proceedings of the National Academy of Sciences of the United States of America 108(35): 14566-14571.
17. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, Figueroa ME, Vasanthakumar A, Patel J, Zhao X, Perna F, Pandey S, Madzo J, Song C, Dai Q, He C, Ibrahim S, Beran M, Zavadil J, Nimer SD, Melnick A, Godley LA, Aifantis I, Levine RL (2011). Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. Cancer Cell 20(1): 11-24.
18. Quivoron C, Couronné L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, Do Cruzeiro M, Delhommeau F, Arnulf B, Stern MH, Godley L, Opolon P, Tilly H, Solary E, Duffourd Y, Dessen P, Merle-Beral H, Nguyen-Khac F, Fontenay M, Vainchenker W, Bastard C, Mercher T, Bernard OA (2011). TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. Cancer Cell 20(1): 25-38.
19. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, **Brudno** Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324(5929): 930-935.
20. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL (2011). Tet-Mediated Formation of

5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science 333(6047): 1303-1307.

21. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science 333(6047): 1300-1303.
22. Chen Q, Chen YB, Bian CJ, Fujiki R, Yu XC (2013). TET2 promotes histone O-GlcNAcylation during gene transcription. Nature 493(7433): 561-564.
23. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473(7347): 343-348.
24. Wu H, D'Alessio AC, Ito S, Xia K, Wang ZB, Cui KR,Zhao K, Sun YE, Zhang Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 473(7347): 389-389.
25. Willemze R, Suciu S, Meloni G, Labar B, Marie JP, Halkes CJ, Muus P, Mistrik M, Amadori S, Specchia G, Fabbiano F, Nobile F, Sborgia M, Camera A, Selleslag DL, Lefrère F Sr, Magro D, Sica S, Cantore N, Beksac M, Berneman Z, Thomas X, Melillo L, Guimaraes JE, Leoni P, Luppi M, Mitra ME, Bron D, Fillet G, Marijt EW, Venditti A, Hagemeijer A, Mancini M, Jansen J, Cilloni D, Meert L, Fazi P, Vignetti M, Trisolini SM, Mandelli F, de Witte T. (2014). High-Dose Cytarabine in Induction Treatment Improves the Outcome of Adult Patients Younger Than Age 46 Years With Acute Myeloid Leukemia: Results of the EORTC-GIMEMA AML-12 Trial. J Clin Oncol. 2014 32(3):219-228
26. Noordermeer SM, Monteferrario D, Sanders MA, Bullinger L, Jansen JH, van der Reijden BA (2012). Improved classification of MLL-AF9-positive acute myeloid leukemia patients based on BRE and EVI1 expression. Blood 119(18): 4335-4337.
27. Noordermeer SM, Tönnissen E, Vissers I, van der Heijden A, van de Locht LT, Deutz-Terlouw PP, Marijt EW, Jansen JH, van der Reijden BA (2011). Rapid identification of IDH1 and IDH2 mutations in acute myeloid leukaemia using high resolution melting curve analysis. British Journal of Haematology 152(4): 493-496.
28. Grossmann V, Tiacci E, Holmes AB, Kohlmann A, Martelli MP, Kern W, et al. (2011). Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. Blood 118(23): 6153-6163.
29. Klein HU, Bartenhagen C, Kohlmann A, Grossmann V, Ruckert C, Haferlach T, et al. (2011). R453Plus1Toolbox: an R/Bioconductor package for analyzing Roche 454 Sequencing data. Bioinformatics 27(8): 1162-1163.
30. Murphy KM, Levis M, Hafez MJ, Geiger T, Cooper LC, Smith BD, et al. (2003). Detection of FLT3 internal tandem duplication and D835 mutations by a multiplex polymerase chain reaction and capillary electrophoresis assay. Journal of Molecular Diagnostics 5(2): 96-102.
31. Verhaak RGW, Goudswaard CS, van Putten W, Bijl MA, Sanders MA, Hugens W, et al. (2005). Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood 106(12): 3747-3754.
32. Roepman R, Letteboer SJF, Arts HH, van Beersum SEC, Lu XR, Krieger E, et al. (2005). Interaction of nephrocystin-4 and RPGRIP1 is disrupted by nephronophthisis or Leber congenital amaurosis-associated mutations. Proceedings of the National Academy of Sciences of the United States of America 102(51): 18520-18525.
33. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. (2010). Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype,

Disrupt TET2 Function, and Impair Hematopoietic Differentiation. Cancer Cell 18(6): 553-567.

34. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature 468(7325): 839-843.
35. Busque L, Patel JP, Figueroa ME, Vasanthakumar A, Provost S, Hamilou Z, et al. (2012). Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nat Genet 44(11): 1179-1181.
36. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, et al. (2012). Clonal Evolution of Preleukemic Hematopoietic Stem Cells Precedes Human Acute Myeloid Leukemia. Sci Transl Med 4(149).
37. Couronne L, Bastard C, Bernard OA (2012). TET2 and DNMT3A Mutations in Human T-Cell Lymphoma. New England Journal of Medicine 366(1): 95-96.
38. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. (2010). DNMT3A mutations in acute myeloid leukemia. N Engl J Med 363(25): 2424-2433.
39. Lu C, Ward PS, Kapoor GS, Rohle D, Turcan S, Abdel-Wahab O, et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature 483(7390): 474-U130.
40. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. (2011). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. Cancer Cell 19(1): 17-30.
41. Li Z, Cai XQ, Cai CL, Wang JP, Zhang WY, Petersen BE, et al. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood 118(17): 4509-4518.

SUPPLEMENTARY INFORMATION

Supplementary methods

Table S1. Primers used for cDNA sequencing of TET2

Exon	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)
3	CAGTTTGCTATGTCTAGGTATTCCGA	TGTGCGTTTTATTCCTCCATTTT
3	CAGAATAGTCGTGTGAGTCCTGAC	GCAATGGAAACACAATCTGGA
3	GAACACACACATGGTGAACTCC	AATTGTGATGGTGGTGGTGG
3	TCCAGGGAACCACAAAGCTAG	GCTTGAGGTGTTCTGACATTGG
3	ACATGTATGCAGCCCTTCTCC	GGGAATCTGCTCTTTGTTGAAA
3	ACCAACATCTCCAGTTCCAA	ATGCACTTGATTTCATGGTCT
3	CAAATGGGACTGGAGGAAGT	GTTTGCTGCTGTTCTTGCTT
3	CAGAAGGACACTCAAAAGCATG	TTGCTGCTCTAAAGCTGGG
3	GAGAATCCACCTGCAAGCTG	TTTCACAAGACACAAGCATCG
3- 6	GAGCAGGTCCTAA	TAAGCTGCTTGGGGATGACC
6- 9	TGTACTACAATGGATGTAAGTTTGCC	GACTTGCCGACAAAGGAAACTAGA
9-10	AAGTCAGGATGTTAGCAGAGCCA	GGCTTTTGAATCAGAATACCCAA
10	CCAGCCCTATGAACTTCTATTC	CCATAACTACAGTGCAGCTCCG
10	CAAACATGGACTATAAAAATGGTGAAC	CACCAGGATCTCCCTCGTCTT
10	GCCGTGGCTCCAACTCAT	GGCAGTGGGGAAAGGTCAC

Table S2. Primers used for Sanger sequencing of DNMT3A exon 7 till 23

Exon	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)
7	CTAATTCCTGGAGAGGTCAAGGTG	AGATGGAGAGAGGAGAGCAGGAC
8	TCTTGCCTCATTCAGATGGAGC	CCTGGGATCAAGAACCTTCCC
9	GTGCTTGCAAGTGTAAGCCTCG	CCTGCACTCCAACTTCCAGG
10	GCCTTCCAGCCTGTCCTGA	CCAACTCTACGGTTCTAGCCAAC
10- 11	GCCTTCCAGCCTGTCCTGA	GTACACCAGCCGCTCTGCA
11- 12	GTTGGCTAGAACCGTAGAGTTGG	CTAAGTGCCTCTGCTACTCTGCC
13	GTTCGAGACTGGGGTCACAGT	AGAAGCGGTGGACACAGTCAG
14	GGTCATGTCTTCAGGGCTTAGG	TGCTACCTGGAATGGAAAGACC
15	TTTCCATTCCAGGTAGCACACC	AGGCTCCTAGACCCACACACC
16	AGGGTGTGTGGGTCTAGGAGC	GCTGTGAAGCTAACCATCATTTCG
17	GTGGGGTAGAATTGTAGCAGGA	ATGAACAAAATGAAAGGAGGCAA
18- 19	TTCCTGTCTGCCTCTGTCCCT	CAGCAGTCCAAGGTAGAAGCCA
20	ATTAACCATTCAGCTCCTCAATCAC	TGCAGTCCCAGCCCACAG
21	GTTATGACGTGTGTGCGTGATT	CCTGCACCGTCTCCTAAATTG
22	GTTTGGCGAGTACCTGGCATATT	GGAAAACAAGTCAGGTGGGAAAG
23	CTGCTGTGTGGTTAGACGGCTT	TGCAATAACCTTCTTGTTTCAGTCA

Supplementary results

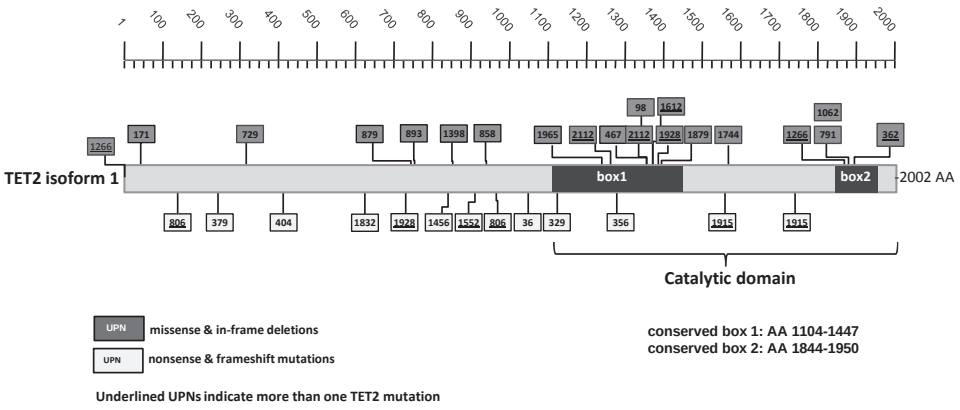


Figure S1. *Schematic representation of the position of TET2 mutations found in 357 patients with AML. A total of 33 mutations was found in 27 patients. Indicated are missense mutations leading to amino acid substitutions (dark boxes) and nonsense and frameshift mutations leading to STOP codons (light boxes). The splice site defect in subject 1266 is shown in the beginning of the protein since it affects the first coding exon of the gene (exon3). The catalytic domain of TET2, including the highly, cross-species conserved domains (box 1 and 2), is indicated. Ruler denotes amino acid number, UPN=unique patient number.*

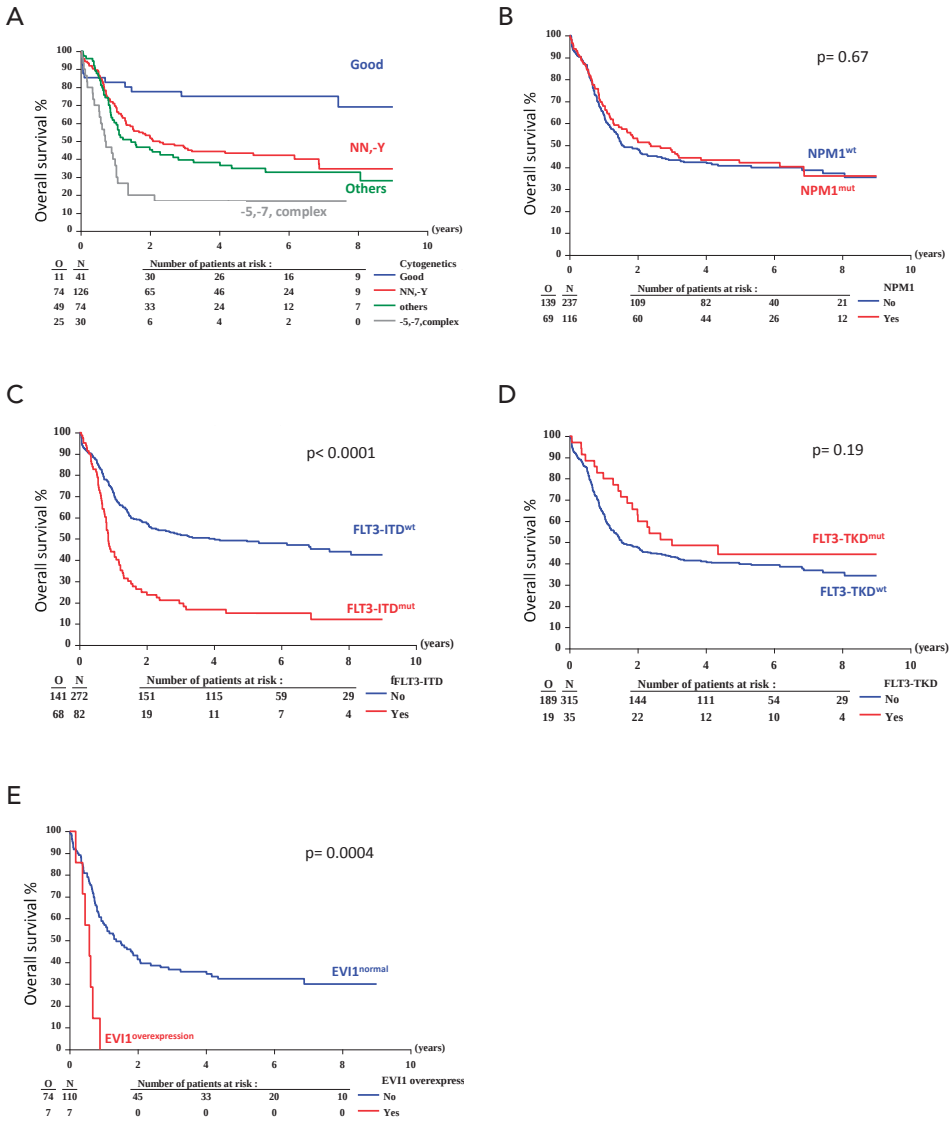


Figure S2. *Univariate analysis indicating the effect of various cytogenetic (A) and molecular markers (B-E) on overall survival in 357 patients with AML. FLT3-ITD and overexpression of EVI1 correlated with significantly inferior overall survival, whereas FLT3-TKD and NPM1 mutations did not significantly affect clinical outcome.*

Chapter 3

Characterization of acute myeloid leukemia based on levels of global hydroxymethylation

Leonie I. Kroeze,* Mariam G. Aslanyan,* Arno van Rooij,
Theresia N. Koorenhof-Scheele, Marion Massop, Thomas Carell,
Jan B. Boezeman, Jean-Pierre Marie, Constantijn J. M. Halkes,
Theo de Witte, Gerwin Huls, Stefan Suciu, Ron A. Wevers,
Bert A. van der Reijden, and Joop H. Jansen,
on behalf of the EORTC Leukemia Group and GIMEMA
*These authors contributed equally to this work

This chapter has been adapted from:
Blood-2013-08-518514

ABSTRACT

Patients with acute myeloid leukemia (AML) frequently harbor mutations in genes involved in the DNA (hydroxy)methylation pathway (*DNMT3A*, *TET2*, *IDH1* and *IDH2*). In this study, we measured 5-hydroxymethylcytosine (5hmC) levels in 206 clinically and molecularly well-characterized younger adult AML patients (≤60 years) included in the EORTC/GIMEMA AML-12 06991 clinical trial, and correlated the 5hmC levels with mutational status and overall survival (OS). In healthy control cells, 5hmC levels were confined to a narrow range (1.5 fold difference), whereas in AML cells, a much wider range was detected (15 fold difference). We identified three 5hmC subpopulations in our patient cohort (low, intermediate and high). The low 5hmC group consisted almost entirely of patients with *TET2* or *IDH* mutations. As expected, *TET2* and *IDH* mutated patients had significantly lower levels of 5hmC compared to patients without mutated *TET2* and *IDH1/2* (both $P<.001$). Interestingly, high 5hmC levels correlated with inferior OS (high vs intermediate 5hmC: $P=.047$, HR=1.81). Multivariate analysis revealed that high 5hmC is an independent poor prognostic indicator for OS (high vs intermediate 5hmC: $P=.01$, HR=2.10).

INTRODUCTION

Alterations in DNA methylation occur in many types of cancer, including hematological malignancies. Global hypomethylation, as well as promoter hypermethylation leading to silencing of various tumor suppressor genes, have been described.[1,2] Since DNA methylation is reversible, it represents a promising target for therapy. In several clinical trials the effect of hypomethylating agents (Azacitidine and Decitabine) on the treatment of patients with myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML) has been assessed,[3-6] which may result in the reactivation of silenced genes. These trials showed promising results for intermediate-1 to high-risk MDS and elderly AML patients by improving complete remission rates, event-free survival and overall survival.

In the last few years our knowledge of DNA (de)methylation has grown extensively. It has been known for a few decades that DNA methyltransferases (DNMT) are responsible for the conversion of cytosine into methylcytosine. DNMT3A and DNMT3B are *de novo* DNA methyltransferases, while DNMT1 is mainly implicated in maintaining methylcytosine marks during DNA replication.[7] DNA demethylation can take place as a passive process when the 5mC mark is not sustained during DNA replication. Only recently, evidence for an active DNA demethylation process has been presented. In 2009 it was discovered that the Ten-Eleven-Translocation (TET) protein family, consisting of TET1, TET2 and TET3, mediates the conversion of 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC).[8] 5hmC can be further oxidized by TET proteins to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which are both recognized by thymine DNA glycosylase (TDG). Subsequently, TDG activates the base excision repair pathway which replaces the modified cytosine with an unmodified cytosine, thereby completing the active demethylation process.[9-11] Neuronal cells and stem cells display remarkably high levels of 5hmC, suggesting that 5hmC may be a rather stable mark with a more versatile function besides being an intermediate in the active DNA demethylation process.[8,12] TET proteins belong to the group of α-ketoglutarate (α-KG) and iron (Fe(II))-dependent dioxygenases. The co-substrate α-KG is produced by the isocitrate dehydrogenase (IDH) proteins, which catalyze the conversion of isocitrate into α-ketoglutarate.[13]

Intriguingly, four genes (*TET2, IDH1, IDH2 and DNMT3A*) important for regulating DNA (hydroxy)methylation are frequently mutated in myeloid malignancies (Figure S1). Loss-of-function mutations in *TET2* have been described in various hematological malignancies,[14-17] including 7-10% of AML patients.[18-20] Heterozygous mutations in *IDH1* and *IDH2* occur in 10-20% of AML patients and are mutually exclusive with *TET2* mutations.[19] Mutant IDH proteins are unable to catalyze the formation of α-KG but gain the function to convert α-KG into the oncometabolite 2-hydroxyglutarate (2HG).[21] *In vitro* studies have shown that 2HG is a competitive inhibitor of the TET proteins.[22,23] This implies that both *TET2* and *IDH1/2* mutations

can impair 5hmC generation, which may explain why mutations in these genes are mutually exclusive. Indeed, AML patients with *TET2* or *IDH* mutations showed decreased 5hmC levels.[24-27] Recently, mutations affecting the methyltransferase *DNMT3A* were described in 20-25% of AML patients.[19,28,29] Interestingly, *DNMT3A* mutations frequently co-occurred with *TET2* or *IDH1/2* mutations,[18] suggesting that an additive effect may result from the combination of these mutations. In most AML studies *DNMT3A* mutations have been associated with a poor prognosis[18,19,28-30] except for a recent study by Gaidzik *et al.*[31] Also *TET2* mutations were associated with a poor prognosis in most studies,[18,19,32-34] whereas the effect of *IDH* mutations is less clear. In several studies *IDH1* mutations correlated with a poor survival in cytogenetically normal AML patients, while the impact of *IDH2* mutations on survival remains inconclusive.[18,35-37]

Several studies have assessed the effect of these mutations on DNA methylation. AML patients harboring *IDH1/2* mutations showed a promoter hypermethylation phenotype and clustered together based on their common DNA methylation profile.[38] The effect of *TET2* mutations on overall DNA methylation remains controversial. In some studies *TET2* mutated patients displayed a hypermethylation signature overlapping with *IDH1/2* mutated patients,[23] while others described no major effect on global methylation or even hypomethylation.[24,39] *In vitro,* most *DNMT3A* mutations led to a decreased capability of the protein to produce 5mC,[29,40] although global 5mC levels in patients with *DNMT3A* mutations were not decreased.[28] Methylated DNA immunoprecipitation (Me-DIP) experiments showed several differentially methylated regions between patients with and without *DNMT3A* mutations.[28,29] No information is available concerning the consequences of *DNMT3A* mutations on 5hmC levels.

The impact of 5hmC levels on survival of AML patients has not been addressed so far. In this study, we measured 5(h)mC levels in a prospectively collected cohort of 206 AML patients included in the EORTC/GIMEMA AML-12 clinical trial, and assessed whether 5hmC levels in AML correlated with clinical and molecular parameters.

MATERIALS AND METHODS

Patient samples

Bone marrow (BM) and peripheral blood (PB) from AML patients (n=206) included in the EORTC/GIMEMA AML-12 06991 clinical trial (NCT0004128, 1999-2009)[41] were obtained after informed consent. To avoid bias, only patients from participating centers where DNA was collected from all consecutive cases were included. In total we obtained DNA from 206 AML patients. After DNA collection, we performed the 5hmC analysis. A consort diagram of the trial and an overview of the patients included and excluded in our study are enclosed in the Supplementary material

(Table S1-2, Figure S2-3). The observed overall survival (OS) rate was slightly lower in the included cohort of patients compared to the non-included group (38.6% vs 41.9% 5-year OS rate, P=.22).The study was conducted in accordance with the Declaration of Helsinki, and was approved by the Institutional Review Board of the Radboud University Nijmegen Medical Centre. The cohort of AML patients (age 16 to 60 years) contained samples of all French-American-British classification subtypes, except for AML M3. The median follow-up at the time of analysis was 6.4 years. PB and BM mononuclear cells, as well as G-CSF-mobilized CD34 positive cells, were obtained from healthy donors (age 18 to 60 years for PB and BM donors and 44 to 67 years for donors of CD34 positive cells).

DNA and RNA isolation

For genomic sequencing and 5(h)mC measurements, DNA was isolated from whole PB or BM of AML patients (at diagnosis or in remission) using salt extraction. We compared the 5hmC values in PB and BM and observed that the levels were comparable (Figure S4). RNA for expression analysis was isolated using RNA-Bee (Tel-Test, inc., Friendswood, TX, USA). DNA from G-CSF-mobilized CD34 positive cells was isolated using salt extraction and DNA from PB and BM mononuclear cells of healthy donors was isolated using the NucleoSpin Blood QuickPure kit (Macherey Nagel, Düren, Germany).

Sequence analysis

Sanger sequencing was performed on PCR-amplified genomic DNA fragments spanning the entire coding region of *TET2* (NM_001127208.2), as previously described.[14] For *IDH1* (ENST00000345146) and *IDH2* (ENST00000330062) the mutational hotspot regions were sequenced. *DNMT3A* exons 7-23 (ENST00000264709) were sequenced using Sanger sequencing or Roche 454-based next generation sequencing (NGS, 454 Life Sciences, Branford, CT, USA). Primers used for *IDH1*, *IDH2* and *DNMT3A* Sanger sequencing can be found in supplementary tables 3 and 4. *DNMT3A* NGS was performed as previously described.[42] In short, this encompasses PCR-based amplicon-library preparation followed by emulsion PCR and pyrosequencing. NGS data were analyzed using R453Plus1Toolbox,[43] GS Variant Analyzer Software 2.6 (454 Life Sciences), and Sequence Pilot version 3.5.2 (JSI Medical Systems, Kippenheim, Germany). For *TET2* and *DNMT3A*, all intron-exon boundaries were included to identify possible splice site mutations. The variations were compared with the dbSNP database to determine whether they represented known SNPs.

TET expression analysis

mRNA levels of *TET2* isoform 1 and isoform 2, *TET1* and *TET3* were determined using quantitative PCR. Specific PCR primers and probes were designed to

discriminate between the different *TET2* isoforms (Table S5). Taqman assays Hs00286756_m1 and Hs00379125_m1 (Life Technologies, Carlsbad, CA, USA) were used to determine *TET1* and *TET3* expression. Quantitative PCR was conducted using an Applied Biosystems Taqman 7900HT Fast Real-Time PCR system (Life Technologies). Expression levels were normalized to *GAPDH* and calculated using the $2^{\Delta CT}$ method.

5mC and 5hmC HPLC-MS/MS measurement

For each patient, 2 µg of DNA (isolated from approximately 1 x 10^6 cells) was degraded into individual nucleosides using DNA degradase plus (Zymo Research, Irvine, CA, USA). The individual nucleosides were measured using a high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) system consisting of an Acquity UPLC (Waters, Milford, MA, USA) containing a Waters Atlantis Hilic column (2.1mm*100mm 3µ) connected to a Micromass Quattro Premier XE (Waters). Quantification was performed using area-based linear regression curves derived from calibration standards containing internal standard solutions. The 5mC and 5hmC levels were calculated as a concentration percentage ratio of %mdC/dG and %hmdC/dG, respectively. The inter-assay coefficient of variation (CV) over a one-year period was 2% for 5mC and 5% for 5hmC (n=49). A detailed description is presented in the supplementary material.

2HG HPLC-MS/MS measurement

A detailed description is presented in the supplementary material.

Bisulfite sequencing

Bisulfite conversion of 500 ng of DNA was performed using the Methyldetector kit (Active motif, Carlsbad, CA, USA), according to the manufacturers protocol. A semi-nested PCR on the converted DNA was carried out to amplify a CpG-rich promoter region in 2 genes frequently hypermethylated in AML; *P15* and *MyoD*. Primer sequences are listed in Table S6. PCR products were cloned into the **pDrive vector** (Qiagen, Valencia, CA, USA), transformed in competent DH12α cells, and plated on a LB agar plate containing 50 µg/ml ampicillin. Per patient, DNA was isolated from 10 colonies and sequenced using Sanger sequencing.

Statistical analysis

5(h)mC differences between two patient groups were statistically tested by performing unpaired T-tests or Mann-Whitney U tests. More than two groups were compared using a One-way ANOVA or Kruskal Wallis test. Pearson and Spearman correlation coefficients were used to calculate correlations between two continuous variables. Two-sided p-values are reported. For negative results that may be scientifically relevant, we determined minimal detectable differences with

a power of 80% (http://biomath.info/power/ttest.htm). To divide the AML patients in subgroups based on their 5hmC level, we fitted a normal mixture model to the data using the mixtools R package.[44] The optimal number of components was determined using the Akaike information criterion (AIC).[45] The results of the OS analysis have been presented in a Kaplan-Meier curve. OS was computed from the date of registration in the EORTC 06991 study until death (by any cause) or until last follow-up (censored observation). The Cox Proportional Hazards Model was used to determine the prognostic importance of each factor analyzed and to obtain hazard ratio (HR) estimates as well as corresponding 95% confidence intervals (95% CI). The following variables were considered: 5hmC levels as continuous or categorical variable, cytogenetic/*FLT3-ITD* risk group (good = inv16, t(8;21); poor = -5/5q-, -7/7q-, complex, *FLT3-ITD*; Intermediate = NN, -Y; others = other cytogenetic abnormalities), white blood cell count (< 25, 25-<100, > 100), % of bone marrow blasts (20-<40%, 40-<60%, 60-<80%, >80%) and age (<30, 30-<40, 40-<50, 50-60 years). Statistical analyses were performed using SAS® (SAS, Cary, NC, USA), R (R Development Core Team, www.r-project.org), and SPSS software (IBM Corporation, Armonk, NY, USA).

RESULTS

Wide distribution of 5hmC levels in AML patients

5hmC levels were measured in 206 AML patients (see **Table 1** for patient characteristics), in healthy purified CD34 positive progenitor cells, as well as healthy BM and PB cells using HPLC-MS/MS. In CD34 positive cells (**Figure 1A**, range 0.037-0.045%; mean 0.041%), and in healthy control cells from BM and PB (**Figure 1B**, range 0.024-0.033%; mean 0.029%) 5hmC levels were tightly clustered, whereas in AML cells a wide range of 5hmC levels was detected (**Figure 1A**, range 0.006-0.091%; mean 0.039%). To determine whether these aberrant 5hmC levels were normalized in remission, we measured 5hmC levels in paired samples at diagnosis and remission from 25 AML patients. For 19/25 patients the type of material (BM or PB) at remission was matched to the material used at diagnosis. In all cases, the remission samples showed 5hmC levels comparable to the levels in healthy BM and PB (**Figure 1B**). In addition, from 5 patients 5hmC was measured at diagnosis, during remission, and at relapse. For all 5 patients, 5hmC levels in remission were comparable to healthy controls (**Figure 1C**). At relapse, 5hmC returned to the levels obtained at diagnosis in all but one patient (UPN128). We found that this patient harbored an *IDH1* mutation at diagnosis which was absent at relapse, which may explain the difference in 5hmC levels between diagnosis and relapse (see next paragraph). For patients with similar 5hmC levels at diagnosis and relapse, the same mutations were present at both time points (**Table S7**).

Table 1. Patient characteristics

	Overall population (n=206)	
	No.	%
Age (years) at diagnosis		
Median (range)	48.0 (16.0-60.0)	
<30	21	10.2
30-<40	36	17.5
40-<50	59	28.6
50-60	90	43.7
Sex		
Male	101	49.0
Female	105	51.0
WHO-ECOG performance status (PS)		
PS 0	122	59.2
PS 1	74	35.9
PS 2	7	3.4
PS 3-4	3	1.5
WBC (x10^9/l) at diagnosis		
Median (range)	24.7 (0.5-312.5)	
< 25	104	50.5
25-99.9	69	33.5
>= 100	33	16.0
FAB subtype		
M0	11	5.3
M1	45	21.8
M2	65	31.6
M4	25	12.1
M5	48	23.3
M6	6	2.9
M7	1	0.5
Unknown	5	2.4
Cytogenetic/*FLT3-ITD* risk group		
Good	24	11.7
Intermediate (NN,-Y)	62	30.1
Poor (-5/5q-,-7/7q-,complex, *FLT3-ITD*)	67	32.5
Others	46	22.3
UNK/ND/Failure	7	3.4

No. indicates number; WHO-ECOG, World Health Organization-Eastern Cooperative Oncology Group; WBC, white blood cell count; FAB, French–American–British classification; UNK, unknown; and ND, not determined.

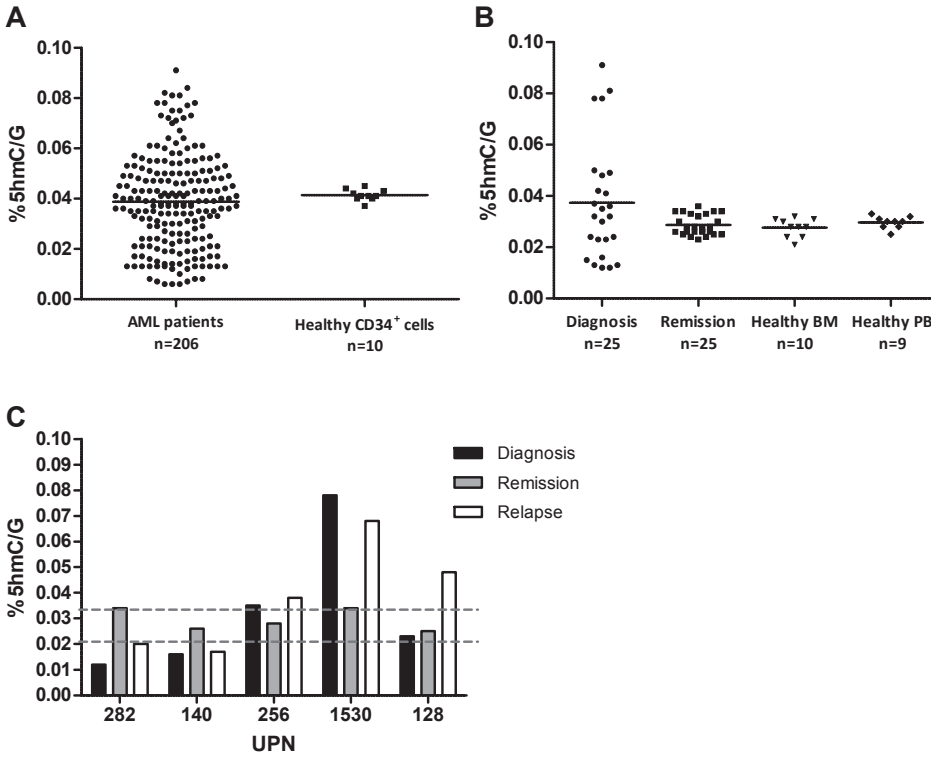


Figure 1. *Aberrant 5hmC levels in AML. (A) In AML cells a broad range of 5hmC levels was detected at diagnosis when compared to healthy CD34 positive cells. (B) Aberrant 5hmC levels at diagnosis were restored to normal levels in remission. (C) For five AML patients 5hmC levels were measured at diagnosis, remission and relapse. The 5hmC levels were in 4 out of 5 cases comparable at diagnosis and relapse. For UPN128 the 5hmC level at relapse was remarkably higher than at diagnosis, due to an IDH1 mutation at diagnosis which was absent at relapse. Dashed lines indicate the range of 5hmC values measured in healthy bone marrow and healthy peripheral blood. UPN indicates unique patient number.*

Decreased global 5hmC levels in patients with TET2 or IDH1/2 mutations

The entire *TET2* gene and the mutational hotspot regions in *IDH1* and *IDH2* were sequenced in 206 AML patients using Sanger sequencing.[18] *TET2* mutations were present in 7% of the patients (n=15), IDH1 R132 in 10% (n=21), IDH2 R140 in 11% (n=22) and IDH2 R172 in 2% (n=4) of the patients. In most cases *TET2*, *IDH1* and *IDH2* mutations were mutually exclusive, except for one patient with an *IDH1* and subclonal *IDH2* mutation (Table S8).

Using HPLC-MS/MS, we confirmed that patients with *IDH* mutations produce 2HG, which is known to inhibit TET proteins (Figure 2A; Table S9).[22,23] Patients with *IDH* mutations had up to 150 times higher plasma 2HG levels compared

to *IDH* wild type patients (*P*=.03). Since a reduced function of TET2, either by mutations or by 2HG-mediated inhibition, may affect both 5mC and 5hmC levels, we compared these levels in patients with and without *TET2* or *IDH* mutations. We observed small, but statistically significant differences in 5mC levels between the different patient groups (Figure 2B). Patients with *IDH* mutations displayed slightly increased 5mC values compared to *TET2/IDH* wild type and *TET2* mutated patients (*P*=.02 and *P*=.001, respectively). In contrast, substantial differences in 5hmC levels were found. Patients with *TET2* mutations had significantly lower 5hmC levels compared to patients without *TET2* or *IDH* mutations (**Figure 2C**, *P*<.001). Patients with two *TET2* aberrations showed a significantly lower 5hmC than patients with a single affected allele (**Figure 2D**, *P*=.003). Furthermore, 5hmC values were decreased in patients with *IDH1/2* mutations compared to *TET2/IDH* wild type patients (**Figure 2C**, *P*<.001). Comparison of the 3 different *IDH* mutational hotspots revealed a slightly lower 5hmC in the subgroup of patients with IDH2 R172 mutations (n=4, **Figure 2D**).

Effect of DNMT3A mutations on global 5mC and 5hmC levels

Since the DNMT3A protein forms 5-methylcytosine, we investigated whether mutations in *DNMT3A* would affect global 5mC and 5hmC levels. *DNMT3A* exons 7-23 were sequenced in 79 patients (**Table S8**). Interestingly, *DNMT3A* mutations frequently co-occurred with *TET2* or *IDH* mutations; 64% of *TET2* mutated and 63% of *IDH* mutated patients harbored a *DNMT3A* mutation. The frequent co-occurrence suggests that an additive oncogenic effect may result from the combination of these mutations. When patients were divided into different subgroups based on co-occurring mutations, mutant *DNMT3A* alone had no strong effect on global 5mC levels (**Figure 2E**). The detected difference between means (DNMT3A/TET2/IDH wildtype vs DNMT3A mutant group) was 0.03, and was not significant (the minimal detectable difference in our cohort was 0.29 %5mC/G with a power of 80%). In addition, *DNMT3A* mutations also did not significantly affect global 5hmC values (**Figure 2F**). The observed difference between means was 0.0056 (the minimal detectable difference with a power of 80% was 0.016 %5hmC/G).

High 5hmC levels confer poor prognosis in AML

In AML a wide range of 5hmC levels was observed when compared to healthy control cells (**Figure 1A**). As 5hmC values were not normally distributed, we used a mixture model analysis to divide the patient cohort in subgroups. Three subpopulations were identified; low 5hmC (0.006-<0.023), intermediate 5hmC (0.023-<0.070) and high 5hmC (0.070-0.091) (see **Table S10** for patient characteristics). We questioned whether 5hmC levels correlated with overall survival by examining the survival of these three groups. Univariate analysis showed that the low and intermediate 5hmC groups had comparable 5-year OS rates (**Figure 3**, 38.3% vs 40.7%). Interestingly,

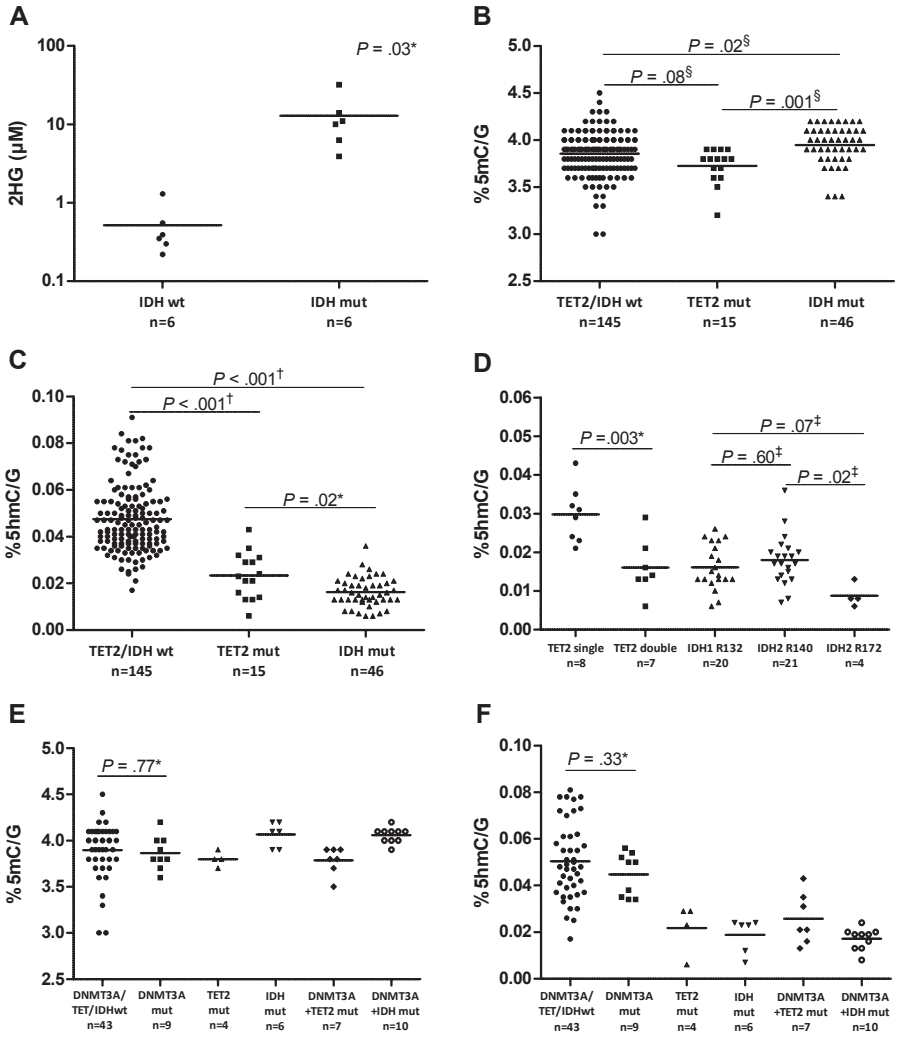
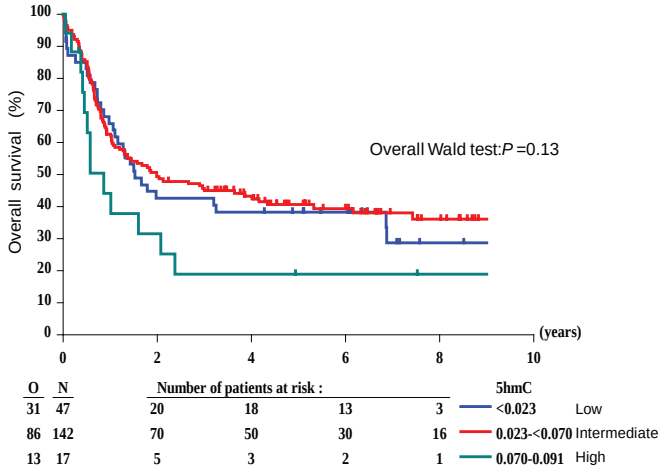


Figure 2. *5hmC levels are decreased in TET2 and IDH mutated patients. (A) Plasma 2HG levels measured by LC-MS/MS were up to 150 times higher in patients with IDH mutations compared to IDH wild type patients. (B) Small differences in 5mC values were present between AML patients. Patients with IDH mutations showed higher 5mC levels than TET2 mutated and TET2/IDH wild type patients. (C) Patients with TET2 or IDH1/2 mutations showed significantly lower 5hmC levels than TET2/IDH wild type patients. (D) Patients with two affected TET2 alleles showed lower 5hmC values than patients with only one affected allele. All types of IDH mutations led to decreased 5hmC levels. The 4 patients with an IDH2 R172 mutation had slightly lower 5hmC values than most patients with an IDH1 R132 or IDH2 R140 mutation. (E) DNMT3A mutations frequently co-occurred with TET2 or IDH mutations. DNMT3A mutations did not significantly affect global 5mC levels. (F) Mutant DNMT3A did also not strongly affect 5hmC values, but the levels were influenced by the presence or absence of accompanying TET2 or IDH mutations. Mut indicates mutated; and wt, wild type. Symbols indicate unpaired T-test (*), Mann-Whitney U test (†), One-way ANOVA (‡) and Kruskal Wallis test (§).*

the high 5hmC group showed a considerably lower 5-year OS rate when compared to the intermediate group (**Figure 3**, 18.9% vs 40.7%) resulting in a significant difference in OS (*P*=.047, HR=1.81, 95%CI=1.01-3.24). A similar trend was observed when event-free survival was analyzed (**Table S11**). Multivariate Cox regression analysis showed that, when adjusted for cytogenetic/*FLT3-ITD* features, high 5hmC levels independently predicted poor overall survival (**Table 2**, high vs intermediate 5hmC: *P*=.01, HR=2.10, 95%CI=1.16-3.80).



Compared groups	p-value	HR	95%CI
low vs intermediate	*P* =.65	1.10	0.73-1.66
high vs intermediate	*P* =.047	1.81	1.01-3.24

Figure 3. *Overall survival according to 5hmC level. Patients were divided into three groups based on a mixture model analysis. Univariate analysis indicates that patients with high 5hmC levels (0.070-0.091) showed a poor overall survival when compared to patients with low (<0.023) and intermediate (0.023-<0.070) 5hmC.*

Correlation of 5hmC levels with clinical and molecular parameters

As shown in figure 2C, low 5hmC levels could in many patients be explained by *TET* or *IDH* mutations. In patients without *TET2* or *IDH* mutations 5hmC levels still varied considerably. Therefore, we examined whether the range of 5hmC values in patients without *TET2* or *IDH* mutations correlated with specific clinical or molecular parameters. There was no significant correlation between 5hmC levels and sex, age, initial performance status, white blood cell count or cytogenetic categories (**Figure S5**). Furthermore, we assessed whether 5hmC levels correlated with the percentage of blasts at diagnosis, since it has been shown that stem cells may contain more 5hmC than differentiated cells.[46] A positive correlation was present

Table 2. Multivariate Cox regression analysis

Parameter	P-value	Hazard ratio	95% Hazard ratio Confidence limits	
5hmC: <0.023 vs 0.023-<0.070	0.44	1.18	0.77	1.81
5hmC: >=0.070 vs 0.023-<0.070	0.01	2.10	1.16	3.80
Cytogenetics/FLT3-ITD: Good vs Intermediate	0.21	0.60	0.25	1.36
Cytogenetics/FLT3-ITD: Others vs Intermediate	0.17	1.44	0.86	2.39
Cytogenetics/FLT3-ITD: Poor vs Intermediate	<0.0001	3.00	1.90	4.68
Cytogenetics/FLT3-ITD: Unknown vs Intermediate	0.02	2.96	1.21	7.21

AIC for this model=1238
Addition of age, treatment arm (data not shown) and/or percentage bone marrow blasts (**Table S12**) in the model did not improve its prognostic discrimination value. Considering 5hmC as a continuous variable in the model (linear and quadratic effect), and adjusting for cytogenetic/FLT3-ITD risk groups provided similar results as the model indicated above (see table S12).

between 5hmC levels and the percentage of blasts in BM (*P*=.003) and PB (*P*=.02) at diagnosis (**Figure S6**). The percentage of BM blasts did however not appear to be of prognostic importance for OS, neither in uni- nor in multivariate analysis. When blast percentage was included in the multivariate model (**Table S12**, model 2), high 5hmC remained an independent poor prognostic factor, thereby showing that 5hmC is not a surrogate marker for percentage of BM blasts. To investigate whether the level of *TET* expression affected the level of 5hmC in AML, we measured mRNA expression of *TET1* (n=31), *TET2* (isoform 1 and 2, n=109) and *TET3* (n=31) in patients without *TET2* or *IDH* mutations. For all *TET* family members, no strong correlation was found with the level of 5hmC (**Figure S7**). Finally, we considered whether 5hmC levels correlated with other specific molecular aberrations. Interestingly, *MLL* aberrations and *AML1-ETO* translocations were enriched in patients with high levels of 5hmC (**Figure 4**). Of note, there was one patient who carried an *IDH2* mutation together with an *MLL* translocation. Of all patients with *MLL* aberrations, this patient showed the lowest value of 5hmC, possibly indicating that the effect of *IDH* mutations (correlating with low 5hmC) overrides the effect of *MLL* translocations (correlating with high 5hmC).

Increased P15 promoter methylation in patients with low 5hmC levels

We examined the methylation status of 2 genes that are frequently hypermethylated in AML (*P15* and *MyoD*)[47] in 10 patients with low and 10 with high 5hmC levels. The majority of patients with low 5hmC (8/10) showed considerable methylation on the *P15* promoter, compared to only 2/10 patients with high 5hmC levels (at least 1 CpG >50% methylated, **Figure 5A**), which suggests that demethylation of the *P15* promoter is altered in most patients with low 5hmC. In contrast, for *MyoD* no difference in promoter methylation was observed between the low and high

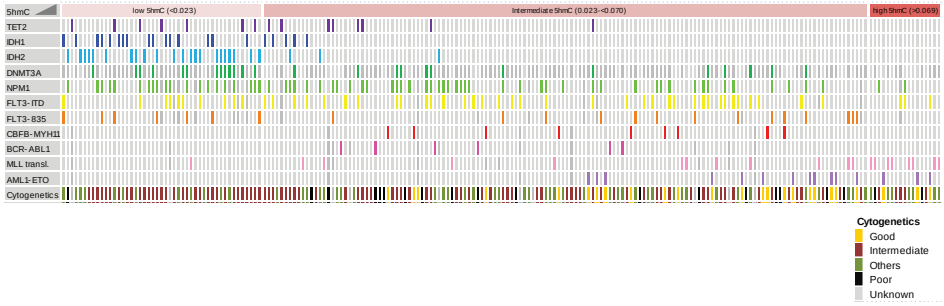


Figure 4. *Molecular markers in different groups of 5hmC levels. Patients are ranked from low to high 5hmC level, and the different 5hmC groups used for the overall survival analysis are indicated. Per patient all mutations are indicated with small colored boxes. Light grey boxes represent no mutation and dark grey boxes not determined/unknown. TET2 and IDH mutations are present in patients with low 5hmC, while AML1-ETO and MLL translocations are more frequently found in patients with high 5hmC levels. ITD indicates internal tandem duplication; inv, inversion; and transl, translocation.*

5hmC groups (**Figure 5B**), indicating that this correlation is promoter specific. Therefore, genome-wide analysis should be performed to determine which genes are differentially methylated between the low and high 5hmC groups.

DISCUSSION

In this study we correlated global 5(h)mC levels with various clinical and molecular parameters in a large prospectively collected cohort of younger adult AML patients (age <=60) included in the EORTC/GIMEMA AML-12 clinical trial. No major differences in global 5mC levels were observed between patients, but 5hmC levels were distributed over a wide range. In remission, 5hmC values were normalized to levels comparable to healthy BM and PB, indicating that the aberrant 5hmC levels at diagnosis were an intrinsic property of the leukemic cells. In line with previous reports, mutations in *TET2* and *IDH1/2* were mutually exclusive,[19,23] and patients with these mutations displayed significantly decreased 5hmC levels.[24-27] Among the patients that did not harbor *TET2* or *IDH* mutations, a wide variation in 5hmC levels was still observed.

Remarkably, *IDH* or *TET2* mutations co-occurred frequently with mutations in *DNMT3A*, indicating that they may co-operate during malignant transformation. The majority of *DNMT3A* mutations are heterozygous missense mutations affecting one specific amino acid (R882), but nonsense and frameshift mutations have also been described.[48] We did not observe significant effects of *DNMT3A* mutations on global 5mC levels, both in patients with mono- and bi-allelic mutations. This suggest that there is still enough remaining DNMT protein (DNMT3B, DNMT1

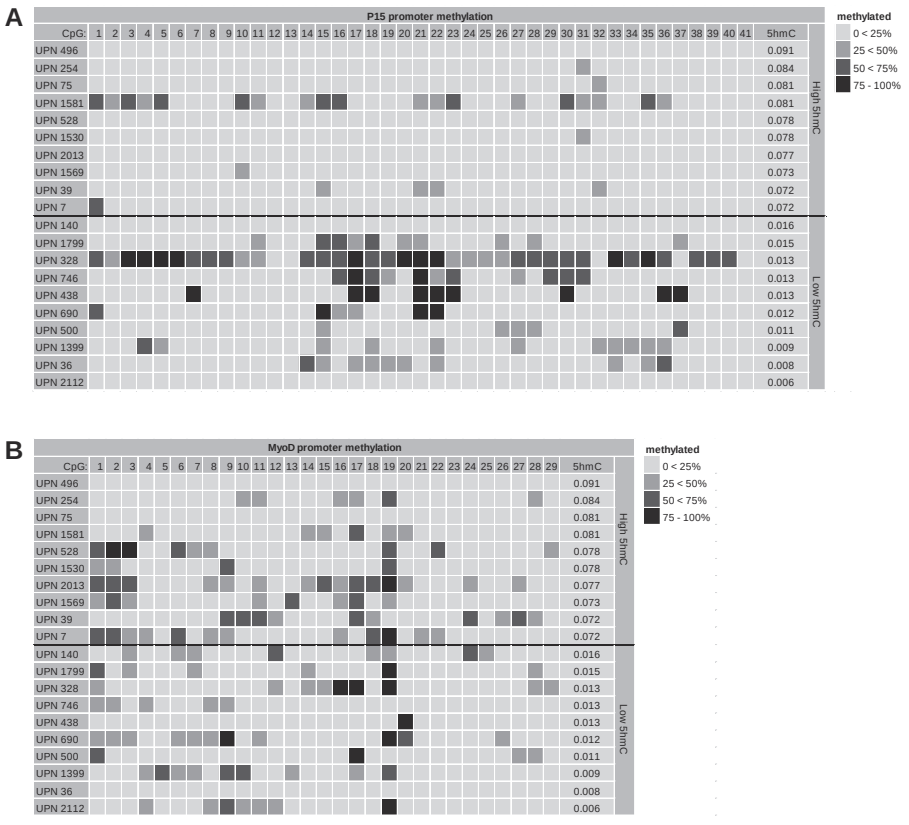


Figure 5. *P15 and MyoD promoter methylation. P15 (A) and MyoD (B) promoter methylation were determined in 10 patients with low and 10 with high 5hmC levels. For each CpG the methylation status is indicated with shaded boxes. Patients with low 5hmC show more P15 promoter methylation than patients with high 5hmC levels, while there is no difference in MyoD promoter methylation. UPN indicates unique patient number.*

and in most cases a remaining *DNMT3A* allele) to form most of the 5mC marks. In addition, we did not find strong correlations between global 5mC levels and survival (Figure S8). This lack of correlation does not necessarily mean that 5mC levels are not relevant, but may rather point to the fact that global DNA methylation patterns obscure local effects at specific genes. In order to identify the important loci for (deregulated) 5mC levels, more targeted strategies, such as MeDIP, should be used.

TET2 and *IDH* mutations both lead to defective conversion of 5mC into 5hmC, which may result in impaired demethylation of DNA, contributing to the malignant transformation of cells. Using bisulfite sequencing, we showed a high level of *P15* promoter methylation in patients with low 5hmC. However, no difference was

observed for the *MyoD* gene, indicating that the correlation is locus-specific. Genome-wide methylation and hydroxymethylation patterns should be determined to see which genes are differentially (hydroxy)methylated between patients with low and high 5hmC levels. Apart from affecting the methylation status of the DNA, other processes might be influenced by 5hmC as well. Several studies showed that the 5hmC mark may attract specific proteins like MECP2 and UHRF2.[49,50] The role of these interactions is not clear yet, but protein complexes that recognize the 5hmC mark might serve as "readers" of the local epigenetic DNA code and induce further chromatin alterations.

After excluding patients with mutations in *TET2* or *IDH*, 5hmC levels still varied considerably in the remaining patients. The cause of this broad range of 5hmC levels could not be attributed to differences in *TET* expression, as no strong correlation was found between *TET1*, *2* or *3* mRNA expression and 5hmC levels. A positive correlation was, however, observed between blast count at diagnosis and 5hmC levels. As it has been shown that stem cells contain more 5hmC than differentiated cells, this might explain the correlation.[46]

We assessed whether 5hmC levels correlated with OS by comparing the survival of the three 5hmC subpopulations present in our cohort (Figure 1A). We showed that high 5hmC levels conferred a poor prognosis in both uni- and multivariate analysis. On the other hand, low 5hmC levels did not affect survival. The low 5hmC group mainly consists of patients with *TET2* or *IDH1/2* mutations (45 out of 47 patients). *TET2* mutations in AML are often associated with a poor prognosis,[32-34] as was also observed in our study.[18] The impact of *IDH* mutations on OS was less pronounced, with *IDH1* mutations showing a trend towards a poor prognosis, whereas *IDH2* mutations had no effect.[18] The differences in OS between these 3 types of mutations indicates that low 5hmC is not the only variable that affects survival in these patients. Interestingly, *AML1-ETO* and *MLL* translocations were more frequently present in patient cells with high 5hmC levels. These findings remain to be confirmed in future studies.

Finally, it would be particularly interesting to investigate how patients with aberrant 5hmC respond to hypomethylating agents, such as decitabine and 5-azacitidine, as these compounds have shown promising results in MDS and elderly AML.[3-6] If a correlation between 5hmC levels and response to these therapeutics could be shown, this would allow a more tailored therapeutic approach.

REFERENCES

1. Roman-Gomez J, Jimenez-Velasco A, Agirre X, et al. Promoter hypermethylation and global hypomethylation are independent epigenetic events in lymphoid leukemogenesis with opposing effects on clinical outcome. *Leukemia*. 2006;20(8):1445-1448.
2. Leone G, Teofili L, Voso MT, Lubbert M. DNA methylation and demethylating drugs in myelodysplastic syndromes and secondary leukemias. *Haematologica*. 2002;87(12):1324-1341.
3. Keating GM. Azacitidine: a review of its use in the management of myelodysplastic syndromes/acute myeloid leukaemia. *Drugs*. 2012;72(8):1111-1136.
4. Joeckel TE, Lubbert M. Clinical results with the DNA hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in patients with myelodysplastic syndromes: an update. *Semin Hematol*. 2012;49(4):330-341.
5. Griffiths EA, Gore SD. Epigenetic therapies in MDS and AML. *Adv Exp Med Biol*. 2013;754:253-283.
6. Ades L, Itzykson R, Fenaux P. Treatment of advanced myelodysplastic syndrome with demethylating agents: azacitidine. *Semin Hematol*. 2012;49(4):323-329.
7. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99(3):247-257.
8. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930-935.
9. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333(6047):1300-1303.
10. He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011;333(6047):1303-1307.
11. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem*. 2011;286(41):35334-35338.
12. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929):929-930.
13. Haselbeck RJ, McAlister-Henn L. Function and expression of yeast mitochondrial NAD- and NADP-specific isocitrate dehydrogenases. *J Biol Chem*. 1993;268(16):12116-12122.
14. Langemeijer SM, Kuiper RP, Berends M, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet*. 2009;41(7):838-842.
15. Delhommeau F, Dupont S, Della Valle V, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med*. 2009;360(22):2289-2301.
16. Kosmider O, Gelsi-Boyer V, Ciudad M, et al. TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. *Haematologica*. 2009;94(12):1676-1681.
17. Langemeijer SM, Jansen JH, Hooijer J, et al. TET2 mutations in childhood leukemia. *Leukemia*. 2011;25(1):189-192.
18. Aslanyan MG, Kroeze LI, Langemeijer SM, et al. Clinical and biological impact of TET2 mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial. *Ann Hematol*. Prepublished on March 29, 2014, as DOI 10.1007/s00277-014-2055-7.
19. Patel JP, Gonen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1079-1089.
20. Gaidzik VI, Paschka P, Spath D, et al. TET2 mutations in acute myeloid leukemia

(AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J Clin Oncol*. 2012;30(12):1350-1357.

21. Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*. 2009;462(7274):739-744.
22. Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell*. 2011;19(1):17-30.
23. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010;18(6):553-567.
24. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010;468(7325):839-843.
25. Konstandin N, Bultmann S, Szwagierczak A, et al. Genomic 5-hydroxymethylcytosine levels correlate with TET2 mutations and a distinct global gene expression pattern in secondary acute myeloid leukemia. *Leukemia*. 2011;25(10):1649-1652.
26. Pronier E, Almire C, Mokrani H, et al. Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. *Blood*. 2011;118(9):2551-2555.
27. Liu X, Guangsen Z, Yi Y, et al. Decreased 5-hydroxymethylcytosine levels are associated with TET2 mutation and unfavorable overall survival in myelodysplastic syndromes. *Leuk Lymphoma*. 2013.
28. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363(25):2424-2433.
29. Yan XJ, Xu J, Gu ZH, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*. 2011;43(4):309-315.
30. Ribeiro AF, Pratcorona M, Erpelinck-Verschueren C, et al. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood*. 2012;119(24):5824-5831.
31. Gaidzik VI, Schlenk RF, Paschka P, et al. Clinical impact of DNMT3A mutations in younger adult patients with acute myeloid leukemia: results of the AML Study Group (AMLSG). *Blood*. 2013;121(23):4769-4777.
32. Weissmann S, Alpermann T, Grossmann V, et al. Landscape of TET2 mutations in acute myeloid leukemia. *Leukemia*. 2012;26(5):934-942.
33. Metzeler KH, Maharry K, Radmacher MD, et al. TET2 mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: a Cancer and Leukemia Group B study. *J Clin Oncol*. 2011;29(10):1373-1381.
34. Chou WC, Chou SC, Liu CY, et al. TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood*. 2011;118(14):3803-3810.
35. Rakheja D, Konoplev S, Medeiros LJ, Chen W. IDH mutations in acute myeloid leukemia. *Hum Pathol*. 2012;43(10):1541-1551.
36. Paschka P, Schlenk RF, Gaidzik VI, et al. IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer adverse prognosis in cytogenetically normal acute myeloid leukemia with NPM1 mutation without FLT3 internal tandem duplication. *J Clin Oncol*. 2010;28(22):3636-3643.
37. Schnittger S, Haferlach C, Ulke M, Alpermann T, Kern W, Haferlach T. IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutated NPM1 status. *Blood*. 2010;116(25):5486-5496.

38. Figueroa ME, Lugthart S, Li Y, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell*. 2010;17(1):13-27.
39. Perez C, Martinez-Calle N, Martin-Subero JI, et al. TET2 mutations are associated with specific 5-methylcytosine and 5-hydroxymethylcytosine profiles in patients with chronic myelomonocytic leukemia. *PLoS One*. 2012;7(2):e31605.
40. Yamashita Y, Yuan J, Suetake I, et al. Array-based genomic resequencing of human leukemia. *Oncogene*. 2010;29(25):3723-3731.
41. Willemze R, Suciu S, Meloni G, et al. High-dose cytarabine in induction treatment improves the outcome of adult patients younger than age 46 years with acute myeloid leukemia: results of the EORTC-GIMEMA AML-12 trial. *J Clin Oncol*. 2014;32(3):219-228.
42. Grossmann V, Tiacci E, Holmes AB, et al. Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. *Blood*. 2011;118(23):6153-6163.
43. Klein HU, Bartenhagen C, Kohlmann A, et al. R453Plus1Toolbox: an R/Bioconductor package for analyzing Roche 454 Sequencing data. *Bioinformatics*. 2011;27(8):1162-1163.
44. Benaglia TC, D.; Hunter, D.R., et al. mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software*. 2009;32(6).
45. Akaike H. Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.). Second international symposium on information theory (pp 267_281). Budapest: Academiai Kiado; 1973.
46. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 2010;38(19):e181.
47. Oki Y, Issa JP. Epigenetic mechanisms in AML - a target for therapy. *Cancer Treat Res*. 2010;145:19-40.
48. Marcucci G, Metzeler KH, Schwind S, et al. Age-related prognostic impact of different types of DNMT3A mutations in adults with primary cytogenetically normal acute myeloid leukemia. *J Clin Oncol*. 2012;30(7):742-750.
49. Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*. 2012;151(7):1417-1430.
50. Spruijt CG, Gnerlich F, Smits AH, et al. Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell*. 2013;152(5):1146-1159.

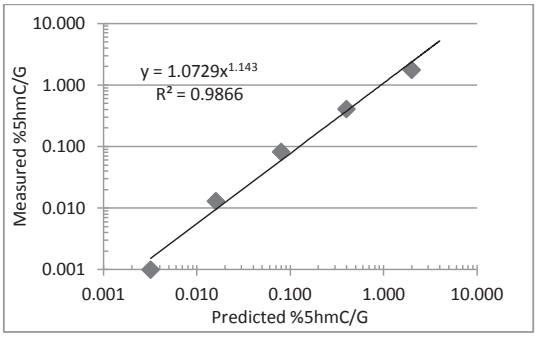
SUPPLEMENTARY INFORMATION

Supplementary materials and methods

5mC and 5hmC HPLC-MS/MS measurement

For each patient, 2 µg of DNA was degraded into individual nucleosides by mixing the DNA with 1x DNA degradase reaction buffer and 5U DNA degradase plus (Zymo Research, Irvine, CA, USA), followed by incubation at 37°C for 5 hours and 70°C for 20 minutes. Hydrolyzed DNA was used for HPLC-MS/MS analysis. Nucleoside calibration standards corresponding to 0.025, 0.05, 0.1, 0.2, 0.5, 1 and 2 µg of DNA, containing 5.3% 5-methyl-2'-deoxycytidine/2'-deoxyguanosine (mdC/dG) and 0.02% 5-hydroxymethyl-2'-deoxycytidine/2'-deoxyguanosine (hmdC/dG), were prepared in deionized water (dH2O). In addition, an internal standard solution was prepared containing 0.02 µM D_2hmdC, 0.27µM $^{13}C_4$-mdC (kindly provided by Prof. Dr. H. Blom) and U-$^{15}N_5$-dG (Cambridge Isotopes Laboratories, Andover, MA, USA) in dH2O. 25 µl dH2O was added to the 25 µl of DNA hydrolysate, and 25 µl 1x degradase reaction buffer to the calibration standards. Subsequently, 25 µl of internal standard solution, 50 µl of methanol (Merck, Darmstadt, Germany) and 500 µl acetonitrile (ACN, Merck) were added to both samples and calibrators. After centrifugation, the eluates were dried under an N_2 flow and dissolved in 75 µl 95% ACN. 5 µl of the calibrators and samples were injected into the HPLC-MS/MS system consisting of an Acquity UPLC (Waters, Milford, MA, USA) containing a Waters Atlantis Hilic column (2.1mm*100mm 3µ) connected to a Micromass Quattro Premier XE (Waters). The mobile phase was made up by ACN and 20 mM ammonium acetate (Sigma-Aldrich, St. Louis, MO, USA) in dH2O. The initial flow rate was set to 400 µl/min. The UPLC flow was directed to the Quattro Premier fitted with an electrospray ionization probe operated in positive mode at unit resolution. The capillary voltage was set at 3 kV. The temperature settings for the source and ion block were 400 ˚C and 100 ˚C respectively. As a drying gas, nitrogen was used at a flow rate of 1100 L/h. The collision cell was operated with argon as the collision gas at a pressure of 0.8 Pa. The cone voltage and collision energy voltage were individually set for the different nucleosides to obtain a maximum response for their deoxyribose loss. Quantification was performed using area-based linear regression curves derived from the prepared calibration standards. The 5mC and 5hmC levels were calculated as a concentration percentage ratio of %mdC/dG and %hmdC/dG, respectively.

In each run, granulocyte DNA, with 5mC and 5hmC levels of 3.9 and 0.027 respectively, was included to monitor inter-assay reproducibility. The inter-assay CV over a one-year period was 2% for 5mC and 5% for 5hmC (n=49). To test the reproducibility of the 5hmC measurements at the extremes, three patient samples with 5hmC values of 0.014 (low), 0.044 (intermediate) and 0.096 (high) were measured 5 times leading to CV values of 5.7%, 2.9% and 1.1%, respectively. The accuracy of the assay was determined by measuring a range of 5hmC values



from 2% to 0.0032% with 5-fold steps, resulting in a R^2 of 0.9866, see graph below. All measurements were performed blinded to clinical and molecular data of the patients.

2HG HPLC-MS/MS measurement

2HG (Sigma-Aldrich) calibration standards of 0.15, 0.3, 0.6, 1.2, 2.5, 5 and 10 µM, and an internal standard solution of 10 µM U-$^{13}C_5$-I-2hg (Chiralex, Nijmegen, the Netherlands) were prepared in dH2O. 100 µl sample was diluted with 50 µl internal standard solution and passed through a Microcon 30 kD ultrafilter (Millipore, Billerica, MA, USA). The filtrate was acidified with 10 µl 4% formic acid (Merck) and 5 µl was injected into the HPLC-MS/MS system consisting of a 232 XL autosampler (Gilson, Middleton, WI, USA) connected to an HP1100 HPLC (Agilent, Palo Alto, CA, USA) with an PFP column (2.1mm*100mm 3µ, Phenomenex, Torrance, CA, USA), and a Micromass Quattro LC (Waters). The mobile phase consisted of dH2O and methanol (Merck) containing 0.3% formic acid (Merck) and was run at a flow rate of 250 µl/min. By running a gradient, 2HG was eluted completely separated from its structural isomers 3-hydroxyglutarate and 2-hydroxy-2-methylsuccinate (the optical D- and L-isomers do still co-elute). The HPLC flow was directed to a Quattro LC tandem mass spectrometer fitted with an electrospray ionization probe operating in the negative mode at unit resolution. The capillary voltage was set at 3 kV, and a cone voltage of 20 V was used. The temperature settings for the source and ion block were 400 ˚C and 100 ˚C respectively. As a drying gas nitrogen was used at a flow rate of 650L/h. The collision cell was operated at 9 eV with argon as the collision gas at a pressure of 0.2 Pa. The mass spectrometer was set to monitor the water loss of both 2HG and U-$^{13}C_5$-I-2hg; recording the mass transitions of m/z 147 to m/z 129 and m/z 152 to 134 m/z, respectively.

Quantification was performed using an area-based linear regression curve derived from the prepared calibration standards containing internal standard solution. The inter-assay (3 months period) CV of a plasma containing 1.0 µM of 2HG was 4% (n=20). All measurements were performed blinded to clinical and molecular data of the patients.

SUPPLEMENTARY TABLES AND FIGURES

Table S1. Patient characteristics

	Patient characteristics AML-12 study	
	Excluded (N=1736)	Included (N=206)
Parameter	N (%)	N (%)
Age (years) at diagnosis		
<30	290 (16.7)	21 (10.2)
30-<40	332 (19.1)	36 (17.5)
40-<50	476 (27.4)	59 (28.6)
50-60	638 (36.8)	90 (43.7)
Sex		
Male	911 (52.5)	101 (49.0)
Female	818 (47.1)	105 (51.0)
Missing	7 (0.4)	0 (0.0)
WHO-ECOG performance status (PS)		
PS 0	805 (46.4)	122 (59.2)
PS 1	740 (42.6)	74 (35.9)
PS 2	160 (9.2)	7 (3.4)
PS 3-4	27 (1.6)	3 (1.5)
Missing	4 (0.2)	0 (0.0)
WBC (x10^9/l) at diagnosis		
< 25	1037 (59.7)	104 (50.5)
25-99.9	499 (28.7)	69 (33.5)
>= 100	191 (11.0)	33 (16.0)
Missing	9 (0.5)	0 (0.0)
FAB subtype		
M0	98 (5.6)	11 (5.3)
M1	435 (25.1)	45 (21.8)
M2	445 (25.6)	65 (31.6)
M3	4 (0.2)	0 (0.0)
M4	346 (19.9)	25 (12.1)
M5	321 (18.5)	48 (23.3)
M6	57 (3.3)	6 (2.9)
M7	3 (0.2)	1 (0.5)
Unknown	27 (1.6)	5 (2.4)
Cytogenetic/*FLT3-ITD* risk group		
Good	165 (9.5)	24 (11.7)
Intermediate (NN,-Y)	578 (33.3)	62 (30.1)

Table S1. Patient characteristics (*Continued*)

	Patient characteristics AML-12 study	
	Excluded (N=1736)	Included (N=206)
Parameter	N (%)	N (%)
Poor (-5/5q-,-7/7q-,complex,*FLT3-ITD*)	144 (8.3)	67 (32.5)
Others	240 (13.8)	46 (22.3)
UNK/ND/Failure	609 (35.1)	7 (3.4)
Treatment randomized		
SD-AraC	859 (49.5)	110 (53.4)
HD-AraC	877 (50.5)	96 (46.6)
SCT in CR		
AutoSCT	484 (27.9)	62 (30.1)
AlloSCT	347 (20.0)	53 (25.7)
Response after induction		
CR	1280 (73.7)	152 (73.8)
NoCR	456 (26.3)	54 (26.2)
EFS status		
NoCR	413 (23.8)	53 (25.7)
CCR	570 (32.8)	61 (29.6)
Relapse	567 (32.7)	69 (33.5)
Death in CR	143 (8.2)	22 (10.7)
Not evaluated	43 (2.5)	1 (0.5)
Survival Status		
Alive	752 (43.3)	76 (36.9)
Dead	984 (56.7)	130 (63.1)

WHO-ECOG indicates World Health Organization-Eastern Cooperative Oncology Group; WBC, white blood cell count; FAB, French–American–British classification; UNK, unknown; ND, not determined; SD, standard dose; HD, high dose; SCT, stem cell transplantation; CR, complete remission; EFS, event free survival; and CCR, continuous complete remission.

Table S2. Duration of survival

Study population	Patients (N)	Observed deaths (O)	Hazard Ratio (95% CI)	P-Value (Log-Rank)	Median (95% CI) (Years)	% at 5 Year(s) (95% CI)
Excluded	1736	984	1	0.2164	2.12 (1.85, 2.48)	41.91 (39.50, 44.30)
Included	206	130	1.12 (0.93, 1.35)		1.68 (1.20, 2.96)	38.57 (31.85, 45.25)

Time used (Years)	Std error	Min AML-12: Duration of survival	Max AML-12: Duration of survival
4.97	1.227	0	10.91
4.35	3.437	0	11.64

95% CI indicates 95% confidence interval.

Table S3. Primers used for *IDH1* and *IDH2* sequencing

Gene position	Forward/Reverse	Primer sequence (5′ to 3′)
IDH1 R132	Forward	tgtaaaacgacggccagtGGCACGGTCTTCAGAGAAGC
	Reverse	caggaaacagctatgaccCAAAATCACATTATTGCCAAGATG
IDH2 R140 and R172	Forward	tgtaaaacgacggccagtCGTCTGGCTGTGTTGTTGCTT
	Reverse	caggaaacagctatgaccAGGATGGCTAGGCGAGGAG
M13 seq primer	Forward	tgtaaaacgacggccagt
	Reverse	caggaaacagctatgacc

M13 sequence primers are indicated by lower-case letters

Table S4. Primers used for sequencing *DNMT3A* exon 7 till 23

Exon	Forward/Reverse	Primer sequence (5′ to 3′)
7	Forward	tgtaaaacgacggccagtCTAATTCCTGGAGAGGTCAAGGTG
	Reverse	caggaaacagctatgaccAGATGGAGAGAGGAGAGCAGGAC
8	Forward	tgtaaaacgacggccagtTCTTGCCTCATTCAGATGGAGC
	Reverse	caggaaacagctatgaccCCTGGGATCAAGAACCTTCCC
9	Forward	tgtaaaacgacggccagtGTGCTTGCAAGTGTAAGCCTCG
	Reverse	caggaaacagctatgaccCCTGCACTCCAACTTCCAGG
10	Forward	tgtaaaacgacggccagtGCCTTCCAGCCTGTCCTGA
	Reverse	caggaaacagctatgaccCCAACTCTACGGTTCTAGCCAAC
10-11	Forward	tgtaaaacgacggccagtGCCTTCCAGCCTGTCCTGA
	Reverse	caggaaacagctatgaccGTACACCAGCCGCTCTGCA
11-12	Forward	tgtaaaacgacggccagtGTTGGCTAGAACCGTAGAGTTGG

Table S4. Primers used for sequencing *DNMT3A* exon 7 till 23 (*Continued*)

Exon	Forward/Reverse	Primer sequence (5′ to 3′)
	Reverse	caggaaacagctatgaccCTAAGTGCCTCTGCTACTCTGCC
13	Forward	tgtaaaacgacggccagtGTTCGAGACTGGGGTCACAGT
	Reverse	caggaaacagctatgaccAGAAGCGGTGGACACAGTCAG
14	Forward	tgtaaaacgacggccagtGGTCATGTCTTCAGGGCTTAGG
	Reverse	caggaaacagctatgaccTGCTACCTGGAATGGAAAGACC
15	Forward	tgtaaaacgacggccagtTTTCCATTCCAGGTAGCACACC
	Reverse	caggaaacagctatgaccAGGCTCCTAGACCCACACACC
16	Forward	tgtaaaacgacggccagtAGGGTGTGTGGGTCTAGGAGC
	Reverse	caggaaacagctatgaccGCTGTGAAGCTAACCATCATTTCG
17	Forward	tgtaaaacgacggccagtGTGGGGTAGAATTGTAGCAGGA
	Reverse	caggaaacagctatgaccATGAACAAAATGAAAGGAGGCAA
18-19	Forward	tgtaaaacgacggccagtTTCCTGTCTGCCTCTGTCCCT
	Reverse	caggaaacagctatgaccCAGCAGTCCAAGGTAGAAGCCA
20	Forward	tgtaaaacgacggccagtATTAACCATTCAGCTCCTCAATCAC
	Reverse	caggaaacagctatgaccTGCAGTCCCAGCCCACAG
21	Forward	tgtaaaacgacggccagtGTTATGACGTGTGTGCGTGATT
	Reverse	caggaaacagctatgaccCCTGCACCGTCTCCTAAATTG
22	Forward	tgtaaaacgacggccagtGTTTGGCGAGTACCTGGCATATT
	Reverse	caggaaacagctatgaccGGAAAACAAGTCAGGTGGGAAAG
23	Forward	tgtaaaacgacggccagtCTGCTGTGTGGTTAGACGGCTT
	Reverse	caggaaacagctatgaccTGCAATAACCTTCTTGTTTCAGTCA
M13 seq primer	Forward	tgtaaaacgacggccagt
	Reverse	caggaaacagctatgacc

M13 sequence primers are indicated by lower-case letters

Table S5. Primers and probes for quantitative PCR

Gene/isoform	Primer/probe	Primer/probe sequence (5′ to 3′)
TET2	Forward primer	AATTTATTGGATACACCTGTCAAGACTC
TET2 isoform 1	Reverse primer	ACCTGCTCCTAGATGGGTATAAAAAG
TET2 isoform 2	Reverse primer	GATAAACGCCATGTGTCTCAGTACA
TET2	Probe	TATGATTTCCCATCTTGC
TET1	Primer/probe mix	Hs00286756_m1 (TaqMan)
TET3	Primer/probe mix	Hs00379125_m1 (TaqMan)

Table S6. Primers bisulfite sequencing

Gene	Forward/Reverse	Primer sequence (5' to 3')
P15	Forward_1	TTTAAAATATTAGTTTTGGTTTTATTGGAT
	Forward_2	TTTTAGTATTTTTGGGTAGGTTTTTT
	Reverse_1/2	ACCTAAACTCAACTTCATTACCCTC
MyoD	Forward_1	GTTTTAAGTTTTTTGTTTTTTG
	Forward_2	GGGATAGAGGAGTATTGAAAGTTAGTTTAG
	Reverse_1/2	AAAACTTCCTCACCCCTAACTTCT
M13 seq primer	Forward	gtaaaacgacggccagt
	Reverse	caggaaacagctatgac

Table S7. Mutations at diagnosis and relapse

UPN	Genetic defects diagnosis	Present at relapse
128	*FLT3-ITD*	yes*
	IDH1	no
140	*FLT3-ITD*	yes
	IDH2	yes
	DNMT3A	yes
256	*FLT3-ITD*	yes
	NPM1	yes
	DNMT3A	yes
282	*FLT3-ITD*	yes
	NPM1	yes
	IDH1	yes
1530	*AML1-ETO*	yes

** Different FLT3-ITD variants at diagnosis and relapse*

Table S8. Mutations in genes involved in DNA (hydroxy)methylation

UPN	*DNMT3A*	*TET2*	*IDH1*	*IDH2*	5mC level	5hmC level
6	N/A	-	p.R132H	-	4.0	0.018
35	p.S714C	-	-	p.R140Q	3.9	0.019
36	p.F640Y	-	-	p.R172L	4.1	0.008
50	p.R882H	-	-	p.R172L	4.0	0.013
80	N/A	-	-	p.R140Q	3.8	0.028
88	p.V563M*	-	-	p.R140Q	4.1	0.019
98	-	p.R1359H	-	-	3.8	0.023
128	-	-	p.R132H	-	4.1	0.023
140	p.R882C	-	-	p.R140Q	4.1	0.016
256	p.R882H	-	-	-	3.7	0.035
265	N/A	-	-	p.R140Q	4.2	0.036
282	N/A	-	p.R132H	-	4.0	0.014
290	N/A	-	p.R132H	-	3.4	0.026
328	N/A	-	p.R132L	-	4.2	0.013
329	p.IVS9-1G>T	p.P1115LfsX2	-	-	3.9	0.035
356	p.R882C	p.N275IfsX18	-	-	3.5	0.021
360	N/A	-	p.R132H	-	3.4	0.013
362	N/A	p.R1896M*	-	-	3.6	0.013
378	N/A	p.R1993W	-	-	3.6	0.024
397	p.R882H	-	p.R132H	-	4.0	0.013
405	p.S714C	-	-	-	4.0	0.050
438	N/A	-	p.R132C	p.R140Q†	4.0	0.013
463	p.R326H	-	-	-	4.2	0.052
484	p.R882C	-	-	p.R140Q	4.1	0.019
566	N/A	-	p.R132C	-	3.7	0.007
573	N/A	-	p.R132C	-	3.8	0.006
582	N/A	-	p.R132C	-	3.4	0.010
596	p.R882H	-	-	-	3.9	0.050
627	p.R882H	-	-	-	3.8	0.034
630	p.G543C	-	-	-	4.0	0.034
671	p.R882H	-	p.R132H	-	4.2	0.024
690	-	-	p.R132C	-	4.2	0.012
746	N/A	-	-	p.R140Q	4.2	0.013
791	-	p.I1873N	-	-	3.9	0.029
806	p.D252X*	p.S137VfsX8	-	-	3.8	0.013
		p.P929LfsX24				
824	N/A	-	p.R132L	-	4.0	0.016
907	N/A	-	p.R132H	-	3.8	0.021
910	N/A	-	-	p.R140Q	3.9	0.014
935	N/A	p.L615AfsX23	-	-	3.2	0.014

Table S8. Mutations in genes involved in DNA (hydroxy)methylation (*Continued*)

UPN	*DNMT3A*	*TET2*	*IDH1*	*IDH2*	5mC level	5hmC level
		p.T229NfsX25				
970	N/A	-	-	p.R140Q	3.8	0.017
977	N/A	-	-	p.R140Q	3.9	0.017
1012	-	-	-	p.R140Q	4.2	0.007
1031	p.R635W	-	-	p.R140Q	4.1	0.020
1079	p.R882H	-	-	-	3.6	0.056
1095	N/A	-	-	p.R140Q	4.1	0.017
1174	N/A	-	-	p.R140Q	3.9	0.012
1176	-	-	p.R132H	-	4.1	0.024
1178	-	-	-	p.R140Q	3.9	0.024
1253	N/A	-	p.R132C	-	3.7	0.013
1266	p.R882H	p.IVS2-2A>G	-	-	3.8	0.016
	p.N501S	p.G1869E				
1269	p.R882H	-	-	p.R140Q	4.0	0.020
1398	p.R882H	p.H850R	-	-	3.9	0.043
1399	N/A	-	-	p.R140Q	4.0	0.013
1433	N/A	-	-	p.R172L	3.9	0.008
1456	p.R882H	p.Q831TfsX15	-	-	3.7	0.031
1546	-	-	p.R132H	-	3.9	0.023
1575	N/A	-	-	p.R140Q	3.9	0.008
1622	N/A	-	p.R132G	-	3.7	0.013
1626	N/A	-	-	p.R140Q	3.9	0.021
1652	p.C517R	-	-	-	3.8	0.054
1663	N/A	-	-	p.R140Q	3.8	0.022
1704	N/A	-	p.R132H	-	3.7	0.019
1763	N/A	-	-	p.R172L	4.1	0.006
1799	N/A	-	p.R132H	-	4.2	0.015
1879	N/A	p.T1397I	-	-	3.8	0.032
1915	-	p.S1758X	-	-	3.7	0.029
		p.T1554SfsX16				
1928	p.IVS9del-18_+1	p.Q749X	-	-	3.9	0.021
		p.D1384N				
2085	p.R882H	-	-	-	3.8	0.038
2111	N/A	-	-	p.R140Q	4.2	0.015
2112	-	p.R1261H	-	-	3.8	0.006
		p.R1359S				

UPN indicates unique patient number; N/A, not analyzed; and - , wild type.
*Homozygous mutation
† Subclonal IDH2 mutation

Table S9. Plasma 2-hydroxyglutarate levels in *IDH* wild type and *IDH* mutant AML patients

Sample	*IDH* mutation	*TET2* mutation	2HG (μM)
1	IDH1 R132H	-	3.9
2	IDH1 R132C	-	10
3	IDH1 R132C	-	14
4	IDH2 R140Q	-	11
5	IDH2 R140Q	-	32
6	IDH1 R132C + subclonal IDH2 R140Q	-	6.3
7	-	Q847X+S128fsX15	0.39
8	-	R1261H+R1359S	1.3
9	-	-	0.22
10	-	-	0.30
11	-	-	0.35
12	-	-	0.55

Table S10. Patient characteristics per 5hmC group

	5hmC groups			Total cohort
	<0.023 (n=47)	0.023-<0.070 (n=142)	0.070-0.091 (n=17)	(n=206)
	No. (%)	No. (%)	No. (%)	No. (%)
Age (years) at diagnosis				
Median (range)	53 (35-60)	46 (16-60)	42 (16-59)	48 (16-60)
<30	0 (0.0)	18 (12.7)	3 (17.6)	21 (10.2)
30-40	1 (2.1)	31 (21.8)	4 (23.5)	36 (17.5)
40-50	14 (29.8)	40 (28.2)	5 (29.4)	59 (28.6)
50-60	32 (68.1)	53 (37.3)	5 (29.4)	90 (43.7)
Sex				
Male	18 (38.3)	77 (54.2)	6 (35.3)	101 (49.0)
Female	29 (61.7)	65 (45.8)	11 (64.7)	105 (51.0)
WHO-ECOG performance status				
PS 0	27 (57.4)	86 (60.6)	9 (52.9)	122 (59.2)
PS 1	19 (40.4)	48 (33.8)	7 (41.2)	74 (35.9)
PS 2	1 (2.1)	5 (3.5)	1 (5.9)	7 (3.4)
PS 3-4	0 (0.0)	3 (2.1)	0 (0.0)	3 (1.5)
WBC (x10^9/l) at diagnosis				
Median (range)	16.9 (0.5-291.0)	25.8 (0.6-312.5)	18.2 (1.0-196.4)	24.7 (0.5-312.5)
< 25	25 (53.2)	69 (48.6)	10 (58.8)	104 (50.5)
25-99.9	11 (23.4)	53 (37.3)	5 (29.4)	69 (33.5)
>= 100	11 (23.4)	20 (14.1)	2 (11.8)	33 (16.0)

Table S10. Patient characteristics per 5hmC group (*Continued*)

	5hmC groups			Total cohort
	<0.023 (n=47)	0.023-<0.070 (n=142)	0.070-0.091 (n=17)	(n=206)
	No. (%)	No. (%)	No. (%)	No. (%)
FAB subtype				
M0	4 (8.5)	6 (4.2)	1 (5.9)	11 (5.3)
M1	14 (29.8)	25 (17.6)	6 (35.3)	45 (21.8)
M2	8 (17.0)	51 (35.9)	6 (35.3)	65 (31.6)
M4	2 (4.3)	22 (15.5)	1 (5.9)	25 (12.1)
M5	17 (36.2)	29 (20.4)	2 (11.8)	48 (23.3)
M6	1 (2.1)	5 (3.5)	0 (0.0)	6 (2.9)
M7	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.5)
Unknown	1 (2.1)	3 (2.1)	1 (5.9)	5 (2.4)
Cytogenetic/*FLT3-ITD* risk group				
Good	0 (0.0)	22 (15.5)	2 (11.8)	24 (11.7)
Intermediate (NN,-Y)	25 (53.2)	33 (23.2)	4 (23.5)	62 (30.1)
Poor (-5/5q-,-7/7q-,complex, *FLT3-ITD*)	11 (23.4)	28 (19.7)	7 (41.2)	46 (22.3)
Others	10 (21.3)	53 (37.3)	4 (23.5)	67 (32.5)
UNK/ND/Failure	1 (2.1)	6 (4.2)	0 (0.0)	7 (3.4)
Percentage BM blasts				
Median (range)	76 (30-98)	64 (20-96)	68 (33-95)	67 (20-98)
< 40%	1 (2.1)	23 (16.2)	2 (11.8)	26 (12.6)
40-<60%	11 (23.4)	35 (24.6)	3 (17.6)	49 (23.8)
60-<80%	12 (25.5)	41 (28.9)	6 (35.3)	59 (28.6)
80-<=100%	23 (48.9)	43 (30.3)	6 (35.3)	72 (35.0)

No. indicates number; WHO-ECOG, World Health Organization-Eastern Cooperative Oncology Group; WBC, white blood cell count; FAB, French–American–British classification; UNK, unknown; ND, not determined; and BM, bone marrow.

Table S11. Event-free survival status according to 5hmC level

			Non-parametric			Cox model	
Parameter	Patients (N)	Observed events (O)	Median (95% CI)(Years)	% at 5 Years (95% CI)	P-value (Log-Rank)	Hazard ratio (95% CI)	P-Value (Score test)
5hmC: <0.023	47	33	0.83 (0.43, 1.54)	31.9 (19.3, 45.3)	0.159 (df=2)	1.00	0.197 (df=2)
5hmC: 0.023-<0.070	141	96	0.92 (0.57, 1.21)	31.4 (23.8, 39.2)		0.96 (0.65, 1.43)	
5hmC: >=0.070	17	15	0.44 (0.00, 1.63)	6.4 (0.4, 25.2)		1.58 (0.85, 2.91)	

95% CI indicates 95% confidence interval

Table S12. Multivariate Cox regression analysis for overall survival

Parameter	P-value	Hazard ratio	95% Hazard ratio Confidence limits	
Model 1 (5hmC as categorical variable, including WBC)*				
5hmC: <0.023 vs 0.023-<0.070	0.18	1.35	0.87	2.10
5hmC: >=0.070 vs 0.023-<0.070	0.004	2.50	1.35	4.17
Cytogenetics/FLT3ITD: Good vs Intermediate	0.26	0.62	0.27	1.43
Cytogenetics/FLT3ITD: Others vs Intermediate	0.13	1.48	0.89	2.47
Cytogenetics/FLT3ITD: Poor vs Intermediate	<0.0001	3.51	2.18	5.65
Cytogenetics/FLT3ITD: Unknown vs Intermediate	0.05	2.96	1.01	6.22
WBC: 25-<100 vs < 25	0.04	1.51	1.01	2.24
WBC: >100 vs < 25	0.17	0.69	0.41	1.17
Model 2 (5hmC as categorical variable, including BM blasts)				
5hmC: <0.023 vs 0.023-<0.070	0.39	1.21	0.79	1.87
5hmC: >=0.070 vs 0.023-<0.070	0.01	2.15	1.18	3.93
Cytogenetics/FLT3-ITD: Good vs Intermediate	0.17	0.56	0.24	1.29
Cytogenetics/FLT3-ITD: Others vs Intermediate	0.15	1.46	0.88	2.44
Cytogenetics/FLT3-ITD: Poor vs Intermediate	<.0001	3.17	2.01	5.03
Cytogenetics/FLT3-ITD: Unknown vs Intermediate	0.01	3.19	1.28	7.99
BM blasts: 40-<60% vs < 40%	0.73	1.12	0.60	2.09
BM blasts: 60-<80% vs < 40%	0.35	1.33	0.74	2.41
BM blasts: 80+ % vs < 40%	0.74	0.91	0.51	1.62
Model 3 (5hmC as continuous variable)				
5hmC x100: Linear effect	0.10	0.42	0.18	0.96
5hmC100 x 5hmC100: Quadratic effect	0.06	1.10	1.01	1.18
Cytogenetics/FLT3ITD: Good vs Intermediate	0.21	0.58	0.25	1.35
Cytogenetics/FLT3ITD: Others vs Intermediate	0.19	1.41	0.84	2.36
Cytogenetics/FLT3ITD: Poor vs Intermediate	<0.0001	2.96	1.89	4.64
Cytogenetics/FLT3ITD: Unknown vs Intermediate	0.02	2.94	1.21	7.14

Model 1: AIC=1234
Model 2: AIC=1241
Model 3: AIC=1240
*WBC was excluded from table 2 (main manuscript) because the HR corresponding to the comparison in OS between WBC 25-<100 and WBC < 25 is around 1.5, whereas the HR for the comparison in OS between WBC >=100 and WBC < 25 is 0.7, which is unexpected as outcome generally decreases with increasing WBC. Since more patients in the WBC>=100 have poor cytogenetic/FLT3-ITD features than in the other WBC groups, such a low HR (+/- 0.7) counterbalances the HR of 3.51 (model 1 above) corresponding to the comparison poor cytogenetic/FLT3-ITD features vs intermediate risk group.

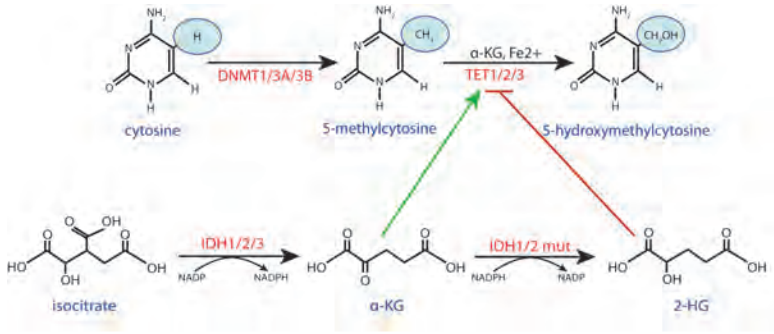


Figure S1. *DNA (hydroxy)methylation pathway. Schematic overview of the DNA (hydroxy) methylation pathway, showing the link between DNMT proteins, TET proteins and wild-type and mutated IDH proteins.*

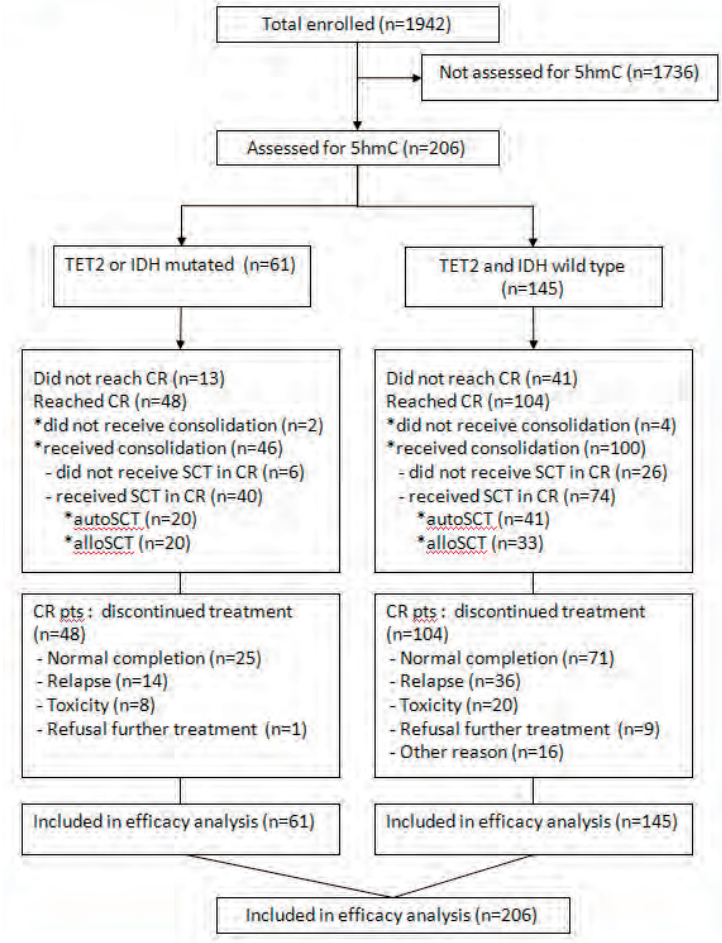


Figure S2. *Consort diagram*

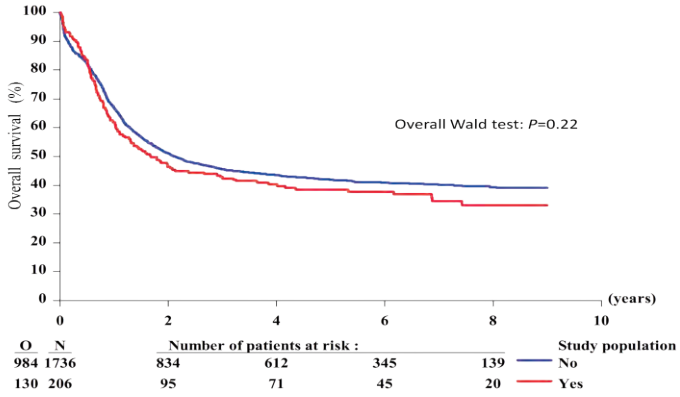


Figure S3. *Overall survival AML-12 cohort. The OS rate for the 206 AML-12 patients included in our study was slightly lower than for the non-included AML-12 patients.*

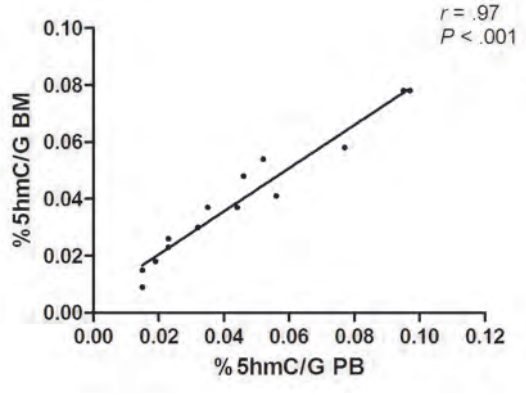


Figure S4. *5hmC levels in peripheral blood and bone marrow are comparable. For 15 patients 5hmC levels were measured in peripheral blood (PB) and bone marrow (BM). These measurements showed comparable levels using both sources. The Pearson correlation coefficient (r) was determined. UPN indicates unique patient number.*

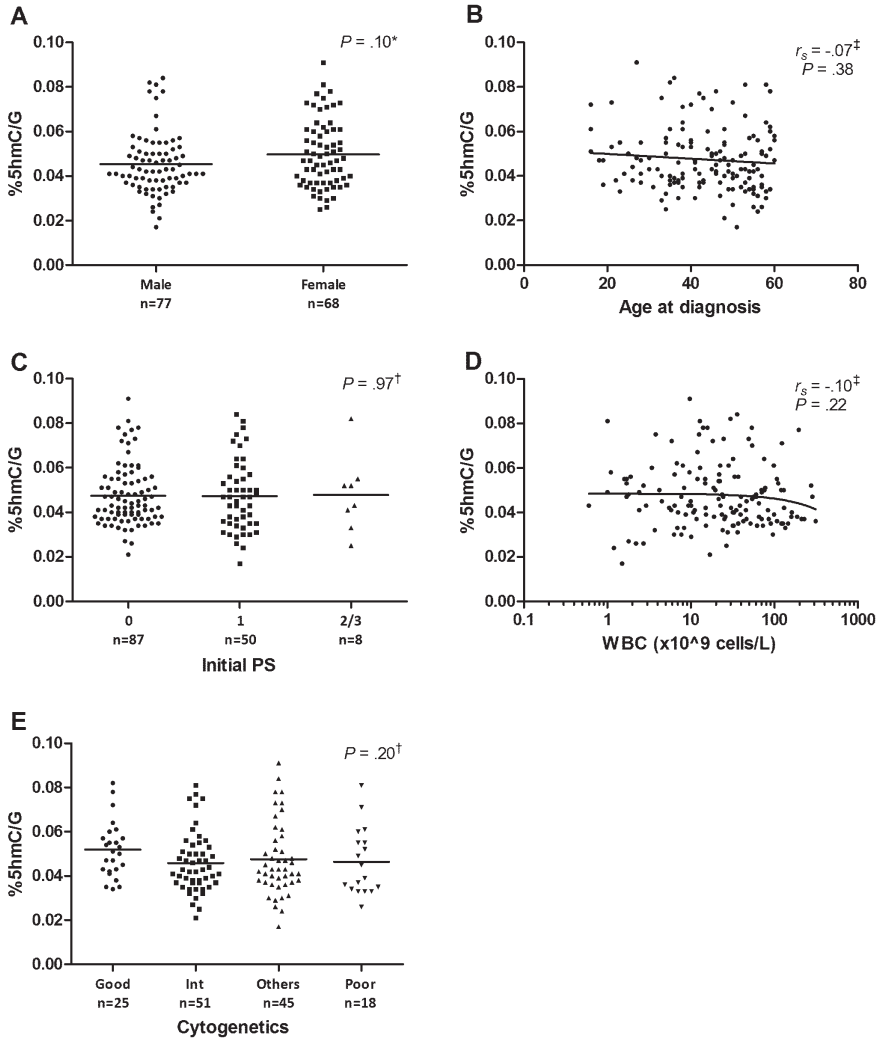


Figure S5. *Correlation of 5hmC levels with clinical and molecular parameters in TET2/IDH wild type patients. No significant correlation was found between 5hmC levels and (A) sex, (B) age, (C) initial performance status (PS), (D) WBC or (E) cytogenetics (FLT3-ITD status not considered). WBC indicates white blood cell count and Int, intermediate. Symbols indicate Mann-Whitney U test (*), Kruskal Wallis test (†), and Spearman correlation coefficient (r_s, ‡).*

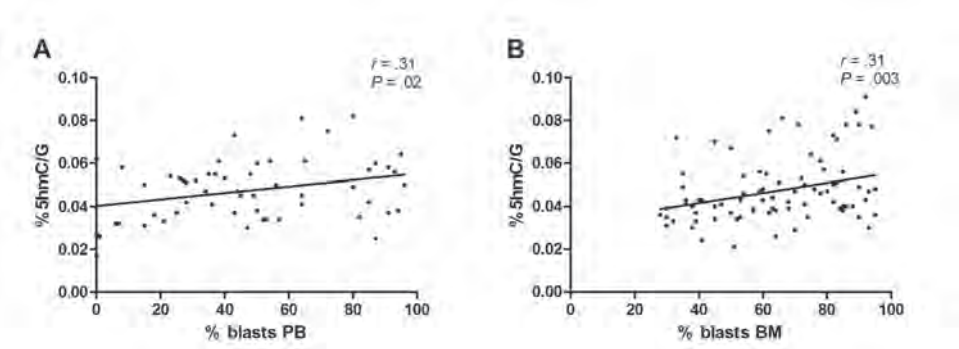


Figure S6. *Blast count correlates with 5hmC levels in TET2/IDH wild type patients. 5hmC values in TET2/IDH wild type patients correlated with the percentage of blasts in (A) peripheral blood (PB) and (B) bone marrow (BM). Pearson correlation coefficients (r) were determined.*

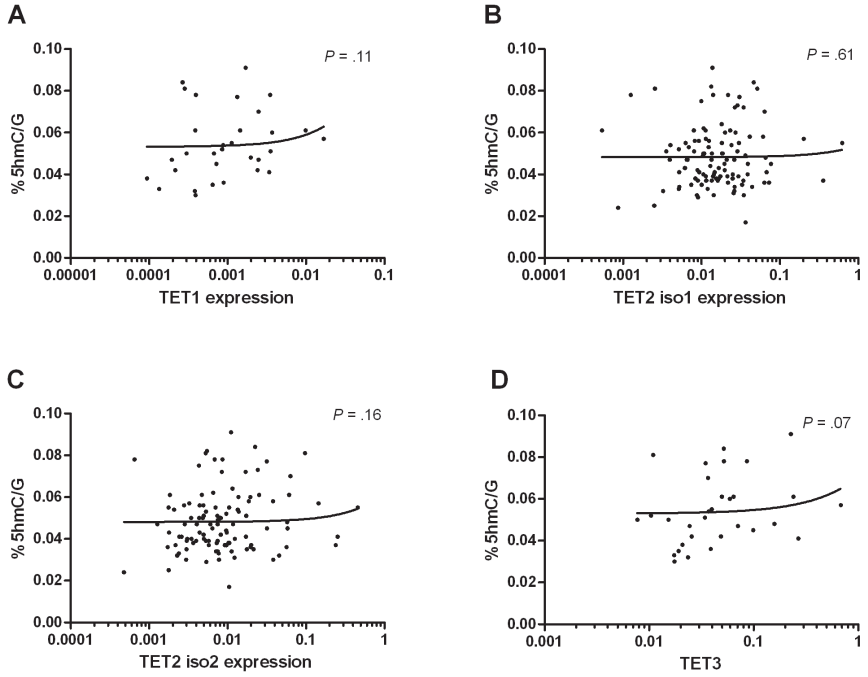


Figure S7. *Correlation between TET expression and global 5hmC values. 5hmC levels in AML patients without TET2 or IDH mutations were correlated with the mRNA levels of (A) TET1, (B) TET2 isoform 1, (C) TET2 isoform 2 and (D) TET3. Expression levels were normalized to GAPDH. No significant correlation could be observed between TET expression and 5hmC levels. Spearman correlation coefficients were determined. Iso indicates isoform.*

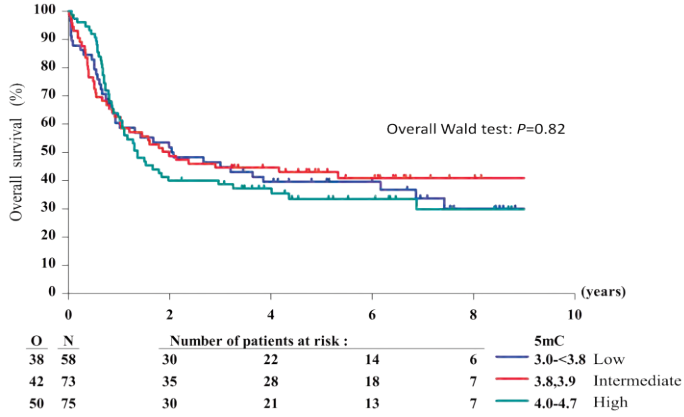


Figure S8. *Overall survival according to 5mC level. No strong correlation could be observed between global 5mC levels and overall survival.*

Chapter 4

Cytosine hydroxymethylation during cell cycle progression

Mariam G Aslanyan [1]*, Leonie I Kroeze [1]*,
Theresia N Koorenhof-Scheele [1]*, Ruth Knops [1], Arno van Rooij [2],
Ron A Wevers [2], Rob Woestenenk [1], Bert A van der Reijden [1],
Joop H Jansen [1]

[1]Dept Laboratory Medicine, Laboratory of Hematology, Radboudumc, Nijmegen, Netherlands
[2]Dept Laboratory Medicine, Laboratory of Genetic, Endocrine and Metabolic Diseases, Radboudumc, Nijmegen, Netherlands

these authors contributed equally to this work

ABSTRACT

Accurate propagation of DNA methylation patterns during replication is fundamental to cells. Recently the biological significance of DNA hydroxymethylation has become evident. It remains, however, unknown whether DNA hydroxymethylation is inherited during cell cycle. To address this question, we investigated global 5-hydroxymethylcytosine levels during various phases of the cell cycle.

Endogenous TET (ten-eleven-translocation) mRNA and protein, as well as global 5-hydroxymethylcytosine levels, were determined in G0/G1, S (early, mid, late), and G2/M phases of the cell cycle. TET mRNA and protein levels did not change. While 5-methylcytosine content remained stable, 5-hydroxymethylcytosine declined to approximately 80% during late S and G2/M phases.

We conclude that, unlike 5-methylcytosine, 5-hydroxymethylcytosine is not sustained during cell cycle progression i.e. the 5hmC mark is not copied immediately during DNA synthesis.

INTRODUCTION

DNA methylation is an epigenetic modification with a crucial role in the regulation of gene expression. It is mediated by the DNA-methyltransferase (DNMT) proteins and represents the covalent transfer of a methyl group to the 5th position of cytosine on DNA. 5-methylcytosine (5mC) is associated with transcriptional repression, playing a key role in processes such as genomic imprinting, X-chromosome inactivation and suppression of repetitive elements and transposons [1]. Germ cell and early embryonic development are also accompanied by dynamic changes in DNA methylation patterns [2]. To ensure their normal development, cells faithfully propagate the 5mC mark during DNA replication through the activity of the DNMT1 protein, the maintenance DNA-methyltransferase. During the S-phase of the cell cycle, DNMT1 is recruited to replication foci, where it preferentially binds hemi-methylated DNA and is thus responsible for copying 5mC onto the daughter strands simultaneously with the ongoing DNA replication [3].

The TET protein family (comprising of TET1, TET2 and TET3) first came into the spotlight when two landmark papers [4, 5] reported frequent TET2 mutations in myelodysplastic syndromes and myeloproliferative neoplasms. Defects in TET2 occur in other myeloid cancers as well [6]. Concurrently, the involvement of the TET proteins in the active DNA demethylation pathway was proven. TET enzymes convert 5mC to 5-hydroxymethylcytosine (5hmC), and further to 5-formyl (5fC) and 5-carboxyl (5caC) cytosine [7-9]. 5fC and 5caC are recognized and excised by the thymine DNA glycosylase (TDG) and replaced by unmethylated cytosines. The TET enzymes have the potential to oxidize all three modifications, but for unclear reasons the process seems to halt at the stage of 5hmC. Compared to 5fC and 5caC, 5hmC is significantly more prevalent in the genome [10]. 5hmC levels are highest in neuronal cells and increase during brain development, suggesting the mark may be more than a mere DNA demethylation intermediate [11, 12]. The existence of 5hmC-specific readers has further stregthened that notion, showing that apart from DNA repair enzymes, a number of transcriptional and chromatin regulators bind 5hmC, as well [13, 14]. In addition, 5hmC levels also have prognostic value in cancer [15, 16]. These data point to the possibility that 5hmC may indeed be a stable epigenetic mark with regulatory functions.

We investigated the levels of all three TET proteins and 5hmC during different phases of the cell cycle, as the question arises whether DNA hydroxymethylation follows a similar pattern of transmission during DNA replication, as does its stable predecessor, 5mC.

MATERIALS AND METHODS

Sorting of cells in different phases of the cell cycle

THP1, OCI-AML3 and HL60 cells were resuspended in fresh medium (37°C) to a concentration of approximately 1*10^6 cells /ml. Hoechst 33342 was added to a final concentration of either 10 µg/mL (OCI-AML3 and HL60), or 5µg/ml (THP1). Subsequently, cells were incubated at 37 °C for 45 minutes. Afterwards, cells were harvested and resuspended in PBS supplied with 1%FCS.

Sorting was performed on a FACS ARIA SORP (BD Biosciences, NJ, USA) using a UV laser and a 450/50 filter. Doublets were discriminated using FSC area vs. width, SSC area vs. width, and Hoechst area vs. height. Sorting was performed based on Hoechst intensity into the following cell populations: G0/G1,S (early, mid and late), and G2/M.

QPCR

THP1, OCI-AML3 and HL60 cells were cultured with 50µg/ml sodium L-ascorbate (Sigma - Aldrich, MO, USA) [17]. Cells from G0/G1, S (early, mid and late) and G2/M phases of the cell cycle were sorted after incubation with Hoechst 33342.

RNA was isolated using a Quick-RNA miniprep kit (Zymo Research, CA, USA). QPCR was performed using the following primer/probe mixes (Applied Biosystems, USA): *TET1* (Hs00286756_m1 FAM-MGB), *TET2* (Hs00289469_m1 FAM-MGB), *TET3* (Hs00379125_m1 FAM-MGB), and β-actin (4310881E VIC-TAMRA).

5mC and 5hmC measurements

Global 5mC and 5hmC measurements were performed using an in-house developed HPLC tandem-MS method (for a detailed description of the entire protocol see [16])

Western blotting

THP1, OCI-AML3 and HL60 cells were cultured with 50µg/ml sodium L-ascorbate (Sigma-Aldrich). Cells from G0/G1, S (early, mid and late) and G2/M phases of the cell cycle were sorted after incubation with Hoechst 33342. In the cases of TET2 and TET3 stainings, cell pellets were lysed in RIPA supplemented with benzonase (Novagen, Merck Millipore, Germany) and Complete Inhibitor of Proteases CIP (Roche, Basel, Switzerland) for 30min. For TET1 staining, nuclear extracts were prepared using the Episeeker nuclear extraction kit (Abcam, Cambridge, UK) according to the manufacturer's instructions. Nuclei were subsequently lysed directly in sample buffer. SDS-PAGE was performed on a 6% PAA gel. Nitrocellulose membranes were used for Western blotting. The following primary antibodies and concentrations were used: TET1 antibody (1:500, Abcam, ab156993); TET2 antibody (1:1000, Diagenode, Liege, Belgium, c15200179); TET3 antibody (1:500, Abcam, ab139321); β-actin (1: 500, Sigma-Aldrich, A5441); cyclin A (1:500, Millipore, MA, USA, D5-375); cyclin B1 (1: 2500, Cell Signaling, MA, USA); GAPDH (1: 5000, Abcam, ab8425).

RESULTS AND DISCUSSION

Sorting strategy

Shortly, cells were incubated with Hoechst 33342 and immediately sorted into 5 sub-populations: G0/G1, S (early, mid, late), G2/M (fig.1A). To ensure that the appropriate sub-populations were selected, cyclin A and cyclin B1 stainings were performed on THP1 cell lysates (fig.1B). As expected, both cyclins were absent in G0/G1. Cyclin A levels were highest in mid and late S phase. Cyclin B1 levels peaked during G2/M phase. Altogether, this proves that cells from different phases of the cell cycle were correctly sorted.

TET mRNA and protein levels remain stable during cell cycle

5hmC is derived from 5mC via the enzymatic activity of the TET proteins. To determine whether TET levels vary during cell cycle, both mRNA and protein levels were determined in THP1, OCI-AML3 and HL60 cells. Three different cell lines were used in order to ensure specific and exclude cell-line specific effects. Q-PCR of *TET1*, *TET2* and *TET3* was performed after sorting, and mRNA levels were normalized to β-actin. No significant changes in *TET* expression were detected (fig.2). In all three cell lines TET3 was most abundant, whereas TET1 levels were lowest. To check whether mRNA levels adequately reflect protein levels, endogenous TET1, TET2

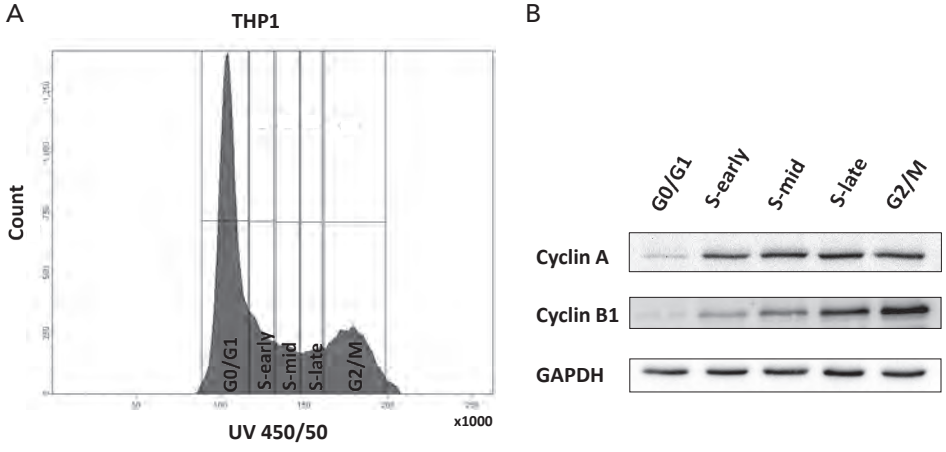


Figure 1. *Sorting strategy (A) FACS plot of THP1 cells. A representative FACS plot of THP1 cells is shown as an example of the sorting strategy we have utilized. Cells from G1, S (early, mid and late), and G2/M were sorted based on Hoechst 33342 intensity; (B) WB stainings of cyclin A and cyclin B1. THP1 cells were sorted into G1, S (early, mid and late), and G2M phases of the cell cycle. WB staining of cyclin A and cyclin B1 showed the expected patterns of expression, with both proteins being absent in G0/G1 and peaking around mid to late S, and G2/M phases. GAPDH was used as a loading control.*

and TET3 protein levels were determined by Western blotting (WB). The protein levels of TET1, TET2 and TET3 remained stable during cell cycle (fig.3).

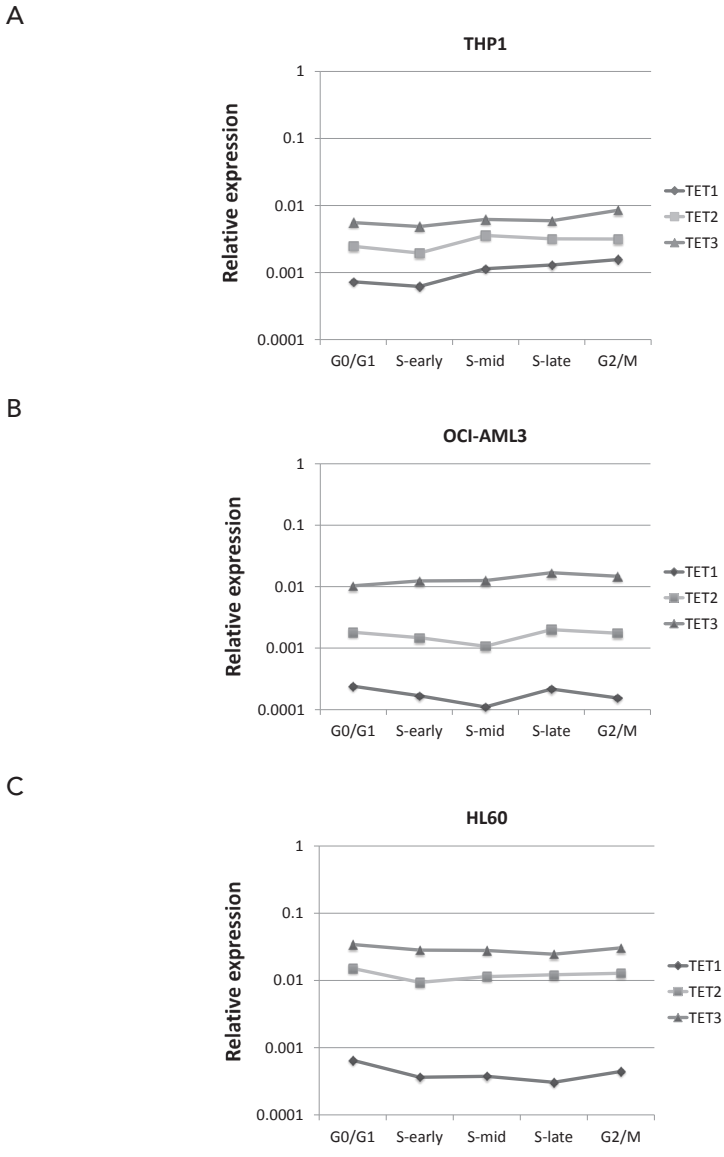


Figure 2. *TET mRNA expression levels do not vary during cell cycle. TET1, TET2 and TET3 mRNA levels were measured in THP1, OCI-AML3 and HL60 (A, B and C) cells in different phases of the cell cycle. Data were normalized to b-actin. There was no significant variation in TET mRNA levels during G0/G1, S (early, mid, late) and G2/M.*

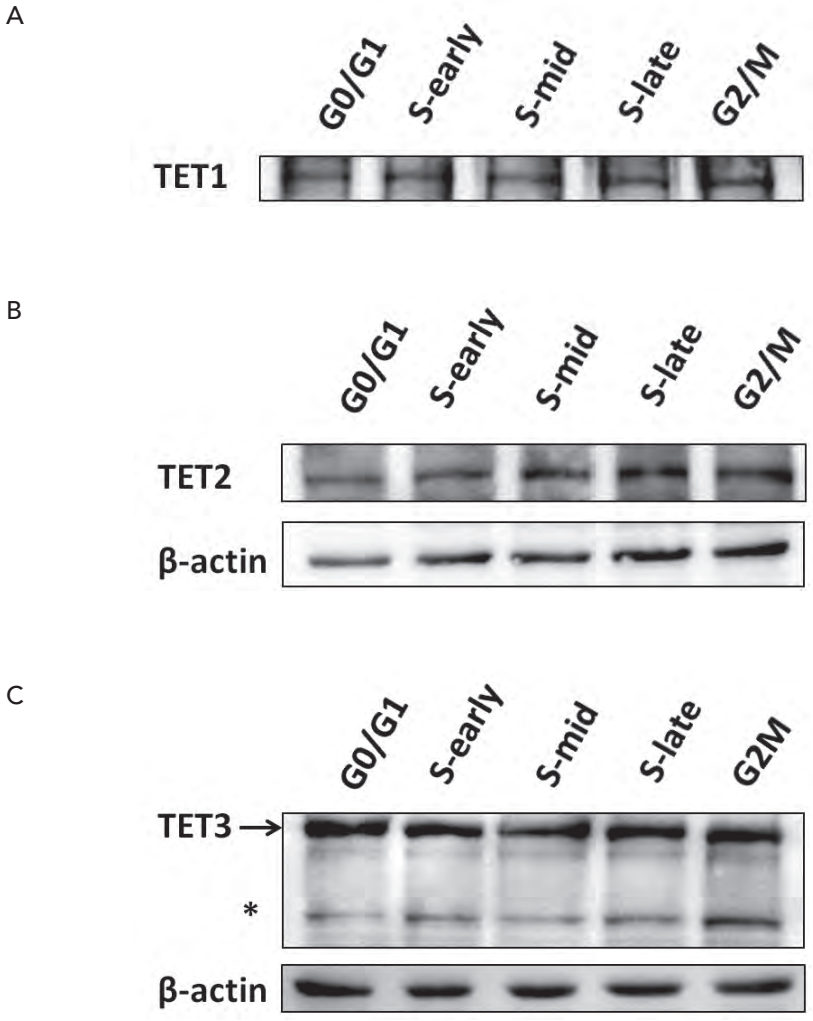


Figure 3. *TET protein levels remain stable during cell cycle. Either nuclear extracts (TET1) (A), or whole-cell lysate (TET2 and TET3) (B and C) from G0/G1, S (early, mid, late) and G2/M phases of the cell cycle, were used for WB. TET1 protein levels were normalized based on equal cell number input. The levels of all three TET proteins remained stable during cell cycle.* *Two different anti-TET3 antibodies consistently recognized this band, which could be a shorter TET3 isoform (ENSP00000307803 according to ensembl database), or a breakdown product. The band indicated with an arrow corresponds to the full-length TET3 protein.

5hmC levels decline during S and G2/M phases of the cell cycle

To determine global levels of DNA (hydroxy)methylation, we made use of an HPLC-MS-MS. In addition, 5mC levels were measured as a control. DNA was isolated from populations of THP1, OCI-AML3 and HL60 cells in G0/G1, S (early, mid and late), and G2/M phases of the cell cycle. 5mC and 5hmC were measured in each population. Each measurement was performed in a biological triplicate (fig.4).

Our results indicate that DNA hydroxymethylation does not follow the same pattern of inheritance as its preceding mark, DNA methylation. Global 5mC levels remained stable during replication since the methyl mark is immediately copied onto the daughter DNA strand, whereas 5hmC levels gradually decreased to 80 -85% during late S and G2/M phases in all three cell lines and were restored to 100% during G0/G1. It would be very interesting to determine whether reinstallment of 5hmC during G0/G1 occurs at the same loci.

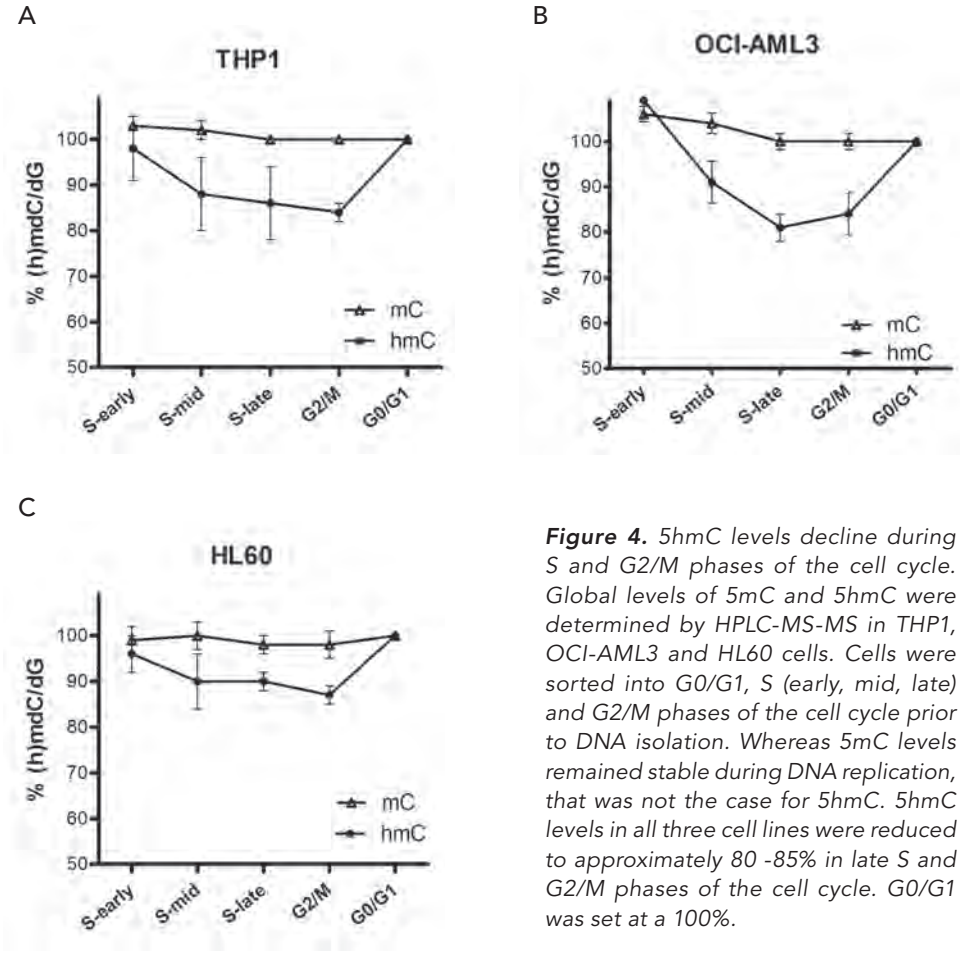


Figure 4. *5hmC levels decline during S and G2/M phases of the cell cycle. Global levels of 5mC and 5hmC were determined by HPLC-MS-MS in THP1, OCI-AML3 and HL60 cells. Cells were sorted into G0/G1, S (early, mid, late) and G2/M phases of the cell cycle prior to DNA isolation. Whereas 5mC levels remained stable during DNA replication, that was not the case for 5hmC. 5hmC levels in all three cell lines were reduced to approximately 80 -85% in late S and G2/M phases of the cell cycle. G0/G1 was set at a 100%.*

5hmC appears to be secondary to the pre-exisiting 5mC mark, i.e no 5hmC can be generated without the prior existence of 5mC. Recently Otani et al, who also showed a decrease of 5hmC levels during DNA replication, demonstrated that the 5mC required by the TET enzymes is generated by the de-novo methyltransferases, DNMT3a and DNMT3b [18]. A possible scenario, therefore, might involve the *de novo* methylation of sites by the DNMT3 proteins, followed by the recruitment of additional factors, which trigger hydroxymethylation. Further studies are required to elucidate the precise mechanism that initiates the formation of 5hmC at specific sites.

Altogether, our data suggest that the 5hmC mark has a different manner of regulation compared to 5mC. This could be logical, as 5mC is essential to a number of fundamental processes, such as genomic imprinting, X-chromosome inactivation and repression of retrotransposons, and this requires a faithful mechanism for maintenance.

We propose that 5hmC, unlike 5mC, is not a template for its own propagation during DNA replication and its formation throughout the cell cycle is dependent on the local epigenetic environment (DNA methylation, as well as histone modifications), rather than a meticulous mechanism for inheritance.

REFERENCES

1. Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes & cancer 2011; 2:607-17.
2. Hackett JA, Surani MA. DNA methylation dynamics during the mammalian life cycle. Philosophical transactions of the Royal Society of London Series B, Biological sciences 2013; 368:20110328.
3. Probst AV, Dunleavy E, Almouzni G. Epigenetic inheritance during the cell cycle. Nature reviews Molecular cell biology 2009; 10:192-206.
4. Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, Stevens-Linders E, van Hoogen P, van Kessel AG, Raymakers RA, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. Nature genetics 2009; 41:838-42.
5. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, et al. Mutation in TET2 in myeloid cancers. The New England journal of medicine 2009; 360:2289-301.
6. Aslanyan MG, Kroeze LI, Langemeijer SM, Koorenhof-Scheele TN, Massop M, van Hoogen P, Stevens-Linders E, van de Locht LT, Tonnissen E, van der Heijden A, et al. Clinical and biological impact of TET2 mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial. Annals of hematology 2014; 93:1401-12.
7. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 2009; 324:930-5.
8. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to

5-formylcytosine and 5-carboxylcytosine. Science 2011; 333:1300-3.

9. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 2011; 333:1303-7.
10. Song CX, Yi C, He C. Mapping recently identified nucleotide variants in the genome and transcriptome. Nature biotechnology 2012; 30:1107-16.
11. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 2009; 324:929-30.
12. Munzel M, Globisch D, Bruckl T, Wagner M, Welzmiller V, Michalakis S, Muller M, Biel M, Carell T. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. Angewandte Chemie 2010; 49:5375-7.
13. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Munzel M, Wagner M, Muller M, Khan F, et al. Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. Cell 2013; 152:1146-59.
14. Iurlaro M, Ficz G, Oxley D, Raiber EA, Bachman M, Booth MJ, Andrews S, Balasubramanian S, Reik W. A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome biology 2013; 14:R119.
15. Lian CG, Xu Y, Ceol C, Wu F, Larson A, Dresser K, Xu W, Tan L, Hu Y, Zhan Q, et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. Cell 2012; 150:1135-46.
16. Kroeze LI, Aslanyan MG, van Rooij A, Koorenhof-Scheele TN, Massop M, Carell T, Boezeman JB, Marie JP, Halkes CJ, de Witte T, et al. Characterization of acute myeloid leukemia based on levels of global hydroxymethylation. Blood 2014; 124:1110-8.
17. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 2013; 500:222-6.
18. Otani J, Kimura H, Sharif J, Endo TA, Mishima Y, Kawakami T, Koseki H, Shirakawa M, Suetake I, Tajima S. Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. PloS one 2013; 8:e82961.

Chapter 5

Conclusion, Discussion & Future directions

1. WHAT THIS THESIS HAS ACHIEVED - CONCLUSIONS

At the commence of this project, in September 2008, little was known about the TET2 gene and protein. Our group had recently identified a mutation in *TET2* in an index MDS patient, and upon sequencing a cohort of MDS patients, discovered that, in fact, TET2 defects are a common event in MDS (25% of the patients). The objective of this thesis was to not only expand upon the clinical involvement of TET2 in AML (a frequent outcome/complication of MDS), but also study the (dys) function of the protein in cells. Chapter 2 of this thesis shows that *TET2* mutations are an independent marker of poor prognosis in AML patients. This emphasizes the importance of testing for *TET2* mutations following a diagnosis of AML, because often the prognosis of a patient based on their genetic background influences the course of treatment. Given the size of the gene, PCR-based sequencing of *TET2* remains cumbersome, but I believe that within the near future next-generation sequencing technologies will prove valuable. In addition, this chapter proves that perhaps most, if not all, TET2 defects result in loss-of-function. TET2 mutants overexpressed in cell line systems failed to produce 5hmC, unlike the wild type protein. As 5hmC levels are influenced by a number of factors, besides the TET proteins, Chapter 3 was dedicated to studying the levels of 5hmC in AML patients and their potential association with patient survival. Interestingly, we identified high 5hmC levels as a marker of poor prognosis.

It remains unclear as to what extent 5hmC can be considered a stable DNA base. An important aspect of 5mC, the precursor of 5hmC, is its manner of transmission during DNA replication. DNA methylation is crucial to the normal development of cells. To ensure its proper distribution through generations, the DNMT1 protein is recruited to replication foci during cell division, where it methylates the newly synthesized DNA strand. In Chapter 4, we reveal that unlike 5mC which is maintained during replication, 5hmC levels decrease during late S and G2/M phases of the cell cycle. This indicates that, likely, DNA hydroxymethylation does not serve as a template for its own propagation, but is rather a consequence of the pre-existing DNA methylation mark.

2. DISCUSSION POINTS AND FUTURE DIRECTIONS

Cooperative effect of genetic mutations affecting DNA methylation and DNA hydroxymethylation in leukemogenesis

An interesting observation was reported in Chapter 2 of this thesis. Upon analyzing co-existence of mutations in AML patients, we noticed that both *TET2* and *IDH1/2* aberrations were frequently accompanied by a mutation in the *DNMT3a* gene ($p= 0.02$ and $p= 0.004$, respectively). This is a somewhat puzzling notion given that these proteins have conflicting functions within the same biological pathway. While

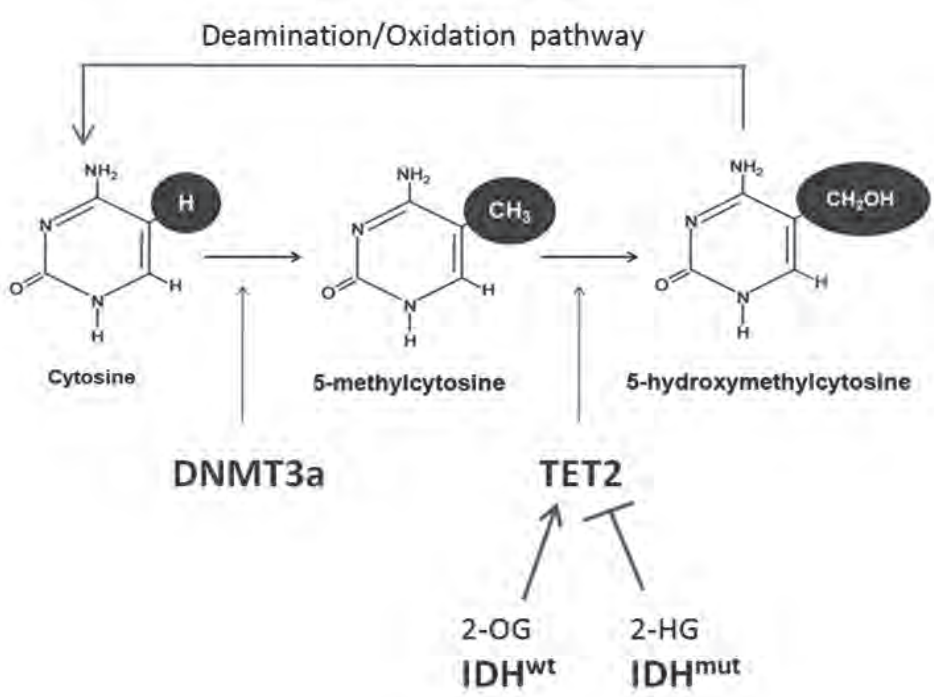


Figure 1. *Leukemia-associated proteins in the DNA methylation and active DNA demethylation pathways*

TET2 and IDH defects perturb DNA hydroxymethylation (i.e. DNA demethylation), DNMT3a aberrations affect the DNA methylation pathway. (fig 1)

Inactivating *TET2* mutations lead to decreased 5hmC levels. *IDH* mutations result in the production of the oncometabolite 2-hydroxyglutarate (2-HG), which also inhibits the TET proteins, thus indirectly contributing to lowering DNA hydroxymethylation levels. Assuming that TET2 and IDH1/2 mutations have, to some extent, overlapping effects, they could be considered a common entity. Therefore, the fact that 67% of AML patients with a *DNMT3a* mutation carry in addition either a *TET2* or and *IDH1/2* defect, cannot be ignored, and is strongly suggestive of a cooperative effect between DNA methylation and DNA hydroxymethylation defects in driving leukemogenic transformation (fig 2).

This hypothesis has been supported by a study published shortly after completion of this thesis. Zhang et al. (Zhang, Combined Effect of *Dnmt3a* Loss-Of-Function and *Idh2* neomorphic mutation Promotes Hematopoietic Malignancy, 2013, ASH abstract) employed a mouse model carrying a double knock-out of the *DNMT3a* gene and an IDH2 R140Q mutation: *DNMT3a (-/-)/ IDH2 R140Q*. Mice developed an MDS/MPN-like disease with frequent progression towards AML. Strikingly, serum

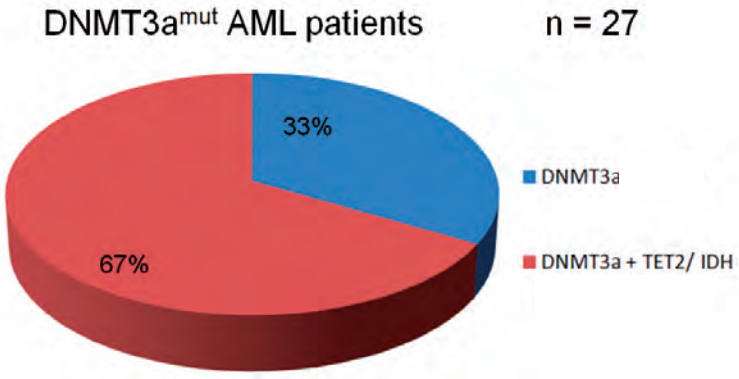


Figure 2. *Frequent co-occurrence of DNMT3a mutations together with TET2 or IDH mutations in AML patients*

from these mice showed levels of 2-HG (2-hydroxyglutarate) which were higher than 2-HG levels measured in mice carrying the IDH2 R140Q mutation alone. This suggests that DNMT3a loss-of-function defects further stimulate the synthesis of 2-HG. It is not yet clear how, but the authors of the study indicate a clear synergistic effect on the metabolome between the two genetic backgrounds. If the DNMT3a protein regulates promoters of genes involved in metabolic processes, that may explain why *DNMT3a* mutations further exacerbate metabolic deregulation, when combined with IDH defects.

Although this study provides an excellent first step towards unraveling the cooperative effect between mutations affecting DNA methylation and DNA hydroxymethylation, a lot remains to be discovered. It would be particularly interesting to create models which bear a DNMT3a mutation together with either a TET2, an IDH1 or an IDH2 defect and compare them to cells carrying the individual mutations, especially in terms of alterations affecting 5mC and its oxidized derivatives (5hmC, 5fC and 5caC). If, indeed, DNMT3a mutations co-occurring with IDH2 R140Q stimulate further metabolic changes, is that also true in the context of IDH1 mutations, and what is the precise mechanism of this cooperation? IDH1/2 mutations induce rather pleiotropic epigenetic alterations, affecting not only DNA (hydroxy)methylation, but the histone code as well. How do DNMT3a and TET2 mutations, on the other hand, contribute to a shared leukemogenic potential? It is reasonable to hypothesize that the combinations of these mutations would have a significant effect on gene expression. Should a particular biological pathway be affected, that would provide a strong rationale for the development of treatments targeting specific subgroups of AML patients carrying *DNMT3a* together with either *TET2* or *DNMT3*a mutations.

5hmC levels in cancer and their association with cancer progression and prognosis

Aberrancies in 5hmC levels may be not only an important factor in cancer initiation/progression, but also affect patient survival. As indicated in Chapter 3, high levels of 5hmC are an independent marker of inferior survival in AML patients. Multiple studies have shown that the prognostic significance of abnormal 5hmC levels extends beyond leukemia and may be a characteristic feature of cancer in general. Yang et al. [1] reported reduced 5hmC levels in human breast, liver, lung, pancreatic and prostate cancer and suggested that decreased TET expression caused the drop in 5hmC levels. Although our own data did not demonstrate a correlation between TET expression and 5hmC levels, we do not exclude the possibility that a reduction of TET levels would be obvious in a particular hematopoietic subtype. The authors propose the implementation of 5hmC levels measurement as a biomarker in cancer. In breast cancer, the leucin zipper LZTS1 had lower levels of 5hmC and TET1 at the promoter region [2]. In addition, this decrease correlated with unfavorable histopathologic parameters, such as lymph node involvement. Furthermore, down-regulation of LZTS1 in breast cancer is associated with metastasis. Jawert et al. [3] showed that TET2 and 5hmC levels are significantly reduced in oral squamous cell carcinoma. Orr et al. [4] reported that decreased 5hmC levels correlated with inferior survival in malignant glioma, and furthermore were associated with tumor progression and aggressiveness. Gambichler et al. [5] also demonstrated reduced 5hmC levels in advanced melanomas. Gastric cancer patients with reduced 5hmC levels had significantly lower survival rate [6]. Low 5hmC also correlated with inferior OS and more aggressive tumor behavior in hepatocellcular carcinoma [7].

The majority of these studies are based on immunohistochemical stainings. If 5hmC levels are to be used as a biomarker in cancer in the future, it is of utmost importance to implement more precise and quantitative methods for 5hmC determination.

Our study showed that most AML cases with low 5hmC could be explained by either a *TET2* or/and *IDH* mutation. The precise contribution of low 5hmC to OS of patients in the context of these mutations needs to be explored in larger patient groups. There were, nonetheless, several unexplained cases in our cohort. It would be interesting to determine what causes the drop of 5hmC in these patients. Several scenarios could be envisaged. In 2013, the group of Rao showed that the CXXC domain protein IDAX (CXXC4) modulates TET2 expression, hence DNA hydroxymethylation [8]. Bioinformatic analysis has shown that IDAX was initially encoded by an ancestral *TET2* gene, which underwent a chromosomal re-arrangement during evolution resulting in gene inversion which split the CXXC domain from the catalytic portion of TET2. IDAX recruits TET2 to DNA and subsequently activates a caspase reaction leading to the degradation, hence downmodulation, of the TET2 protein. In addition, it has been previously

demonstrated that IDAX is an inhibitor of the Wnt signalling pathway and has been implicated in malignant renal cell carcinoma and colonic villous adenoma [9]. Therefore, it is likely that mutations or de-regulated expression of IDAX could also contribute to disturbances in 5hmC levels. In the past few years several AML genomes have been sequenced. As no study thus far has reported putative mutations of *IDAX* (CXXC4) in AML, and decreased expression of the gene has previously been associated with renal carcinoma, the expression levels of IDAX should be investigated in the context of AML.

Alternatively, TET2/5hmC levels may be downregulated by miRNAs. microRNAs are small non-coding RNAs which bind to 3'UTRs of mRNA molecules and infuence their stability. Up to date, several miRNA have been shown to affect TET2 levels. miRNA- 29a levels inversely correlate with TET levels [10]. mi-RNA22 is upregulated in MDS and leukemia and its expression confers poor OS of patients. TET2 is key target of mi-RNA22 [11]. mi-RNA22 downregulates TET2 expression in patients, hence decreases 5hmC levels. Most recently, Cheng et al. [12] reported that mi-RNA125b, mi-RNA29b, mi-RNA29c, mi-RNA101 and mi-RNA7 target TET2, and are exclusively overexpressed in TET2 wild type AML.

Despite the afore-mentioned hypotheses, the possibility remains that some case will remain unexplained. The best way to approach those would be the utilization of next-generation sequencing technologies.

Not only low, but also high 5hmC levels, affect patient outcome. Chapter 3 of this thesis shows that high 5hmC levels are an independent indicator of poor prognosis in AML. In addition, AML1-ETO and MLL translocations were found more frequently in patients with high 5hmC. Huang et al. [13] established a link between TET1 and MLL demonstrating that TET1 is a direct target of MLL-fusion proteins. Furthermore, TET1 was found to be up-regulated in MLL-rearranged leukemias, resulting in a global increase of 5hmC levels. Generally, the AML-ETO translocation is associated with relatively good prognosis, unless c-kit mutations are present. It would therefore be very valuable to determine whether high 5hmC occurs in patients carrying AML-ETO in combination with c-kit abnormalities.

Based on the fact that TET mutations alone are unable to cause leukemia, the group of Helin (Rasmussen, TET Proteins and Histone Modifications in Hematopoiesis, 2013, ASH abstract) set out to determine which additional aberrations might cooperate with the TETs in the initiation of leukemia. The study suggests that AML1-ETO fusion together with TET2 mutations synergize causing an aggressive AML-like disease in mice. DNA methylation was indeed altered displaying progressive hypermethylation of gene regulatory elements. However, the authors make no mention of global 5hmC levels in mice carrying both defects. Moreover, in patients, TET2 mutations and AML1-ETO co-occur only rarely. It remains therefore unclear how, and whether, the AML1-ETO fusion protein alters DNA hydroxymethylation levels. ETO is known to interact with HDACs and

co-repressors, such as Sin3a [14, 15], and the fusion protein can recruit DNMT1 to target genes [16]. The substantial involvement of AML1-ETO in epigenetic processes gives a strong rationale to investigate in-depth the link between high 5hmC and AML1-ETO fusion.

Since AML1-ETO, in the absence of secondary events [17] is also insufficient for leukemogenesis, possibly additional factor(s) contribute to the increase of 5hmC levels in AML patients.

5hmC – the "sixth base"?

Despite our growing knowledge of DNA (de)methylation and the proteins involved in these epigenetic processes, the precise mechanisms that govern and regulate them, need to be defined. The discovery of the TET proteins and their role in the formation of 5hmC marked a new era in the area of DNA demethylation - a pathway that had remained controversial for many years. The TET proteins oxidize 5mC consecutively to 5hmC, 5fC and finally 5caC; 5fC and 5caC are recognized and excised by the TDG protein and replaced by an unmethylated C. This process could have been construed as simple and straightforward, were it not for the puzzling fact that the TET proteins seem to halt at the stage of 5hmC. The explanation for this is presently unknown and has evoked discussion as to the function(s) of DNA hydroxymethylation. Initially, researchers expected that as a DNA demethylation intermediate 5hmC would simply counteract the consequences of 5mC and reverse gene silencing. Surprisingly, the group of Helin showed that a number of Tet1-bound genes were downregulated. This implies that 5hmC does nor per se indicate activation of gene expression and may be subject to more complex regulation than anticipated [18]. Moreover, a recent study claims that despite their overlapping functions, the TET proteins in fact have different roles in the demethylation cascade, with TET2 and TET3 being primarily responsible for the further oxidation of 5hmC into 5fC and 5caC [19]. Based on this study, one could hypothesize that while TET1 is responsible for a poised state i.e. a halt at the level of 5hmC, TET2 and TET3 may indeed be associated with upregulation of gene expression.

The presence of 5hmC varies across different tissues and cell types. Highest levels of 5hmC have been measured in the brain, especially in the brain cortex [20]. This is an extremely interesting observation, not only in terms of 5hmC function, but also evolution of the brain. The cerebral cortex is associated with higher brain function - cognition, reasoning, decision-making, problem-solving, visual processing to name a few. One can envisage that such complex processes require an adaptable cellular machinery, one equipped to quickly translate input into gene expression. In this context, 5hmC, even 5fC and 5caC, could be viewed as an additional layer to the complexity of DNA methylation regulation in mammals [21] – a step in the evolutionary process. It would therefore be very important to determine how the distribution of the TET proteins and 5hmC has changed throughout evolution.

While 5hmC levels are high, 5fC and 5caC have not been found to accumulate in brain tissues [22]. They are present in much lower levels in the genome and are believed to be only transient, short-lived intermediates in the DNA methylation cascade. 5hmC, but also 5fC, attracts specific "readers" [23, 24]. Those include DNA repair enzymes, but also transcriptional and chromatin regulators.

Altogether the data discussed so far neither prove, nor invalidate, the premise that 5hmC is a stable epigenetic mark with functions beyond DNA demethylation.

In Chapter 3 we hypothesized that based on its lack of a clear inheritance pattern during cell cycle, 5hmC should not be viewed as a stand-alone mark. Global 5mC levels remained stable during DNA replication, since the methyl mark is immediately copied onto the daughter DNA strand by DNMT1. However, 5hmC levels declined to 80 -85% during late S and G2/M phases in all three cell lines that we tested. Albeit not seemingly dramatic, the consequences of this decrease should be considered in the context of the asymmetric distribution of 5hmC. Following base-resolution sequencing of 5hmC in the mammalian genome, Yu et al. [25] showed that although ~92% of 5mC was symmetrically modified, that was true for only 21% of 5hmC sites. The major implication of this finding would be the complete loss of 5hmC on a number of daughter DNA strands post replication. Nonetheless, according to our data 5hmC levels are restored in G0/G1. It remains unclear whether 5hmC is re-established at the same genomic positions, or deposited at different sites. It would be particularly interesting to gain more insight into DNA hydroxymethylation changes occurring both globally and at local promoters during various phases of the cell cycle, using base-resolution sequencing technologies.

5hmC is dependent on 5mC generated by the de-novo methyltransferases, DNMT3a and DNMT3b [26]. DNMT3a and DNMT3b, unlike DNMT1, can recognize and methylate cytosine in the context of hemi-hydoxymethylated DNA [27], suggesting that 5hmC could by itself trigger local epigenetic changes. At sites where 5hmC is completely lost, however, there must be other factors responsible for restoring the mark at these positions. IDAX has been shown to regulate the activity of the TET proteins [8]. Direct interactions exist between TET1 and EZH2, a member of the PRC2 complex involved in histone methylation [28]. Nevertheless, the first step in generating 5hmC is the recruitment of the DNMT3 proteins and the establishment of the DNA methylation signature. Although not a definitive proof, our study clearly supports the notion that 5hmC, and likely 5fC and 5caC, add to the complexity of DNA methylation regulation, but do not rank equal with it. The oxidized derivatives of 5mC could be merely intermediates in the DNA deymethylation pathway. Unravelling the intricate mechanisms that govern this pathway remain a fascinating challenge that will certainly inspire a number of researchers in the foreseeable future.

3. CONCLUDING REMARKS

Undeniably research, much like fashion, is driven by trends. There is no doubt that the current "sweethearts", the TET proteins and 5hmC, have taken the limelight and attracted the attention of scientist from both the hematology and the epigenetics fields, and are slowly expanding their influence to include cancer research and neurobiology, among others. This has resulted in a number of crucial discoveries in a short time. It is my belief, that the following years will witness significant progress on several fronts. Firstly, scientists need to unveil the precise mechanism governing the DNA demethylation cascade. Despite their seemingly overlapping functions, it is clear that the TET proteins likely have unique activities too [21]. Establishing what those are will thus contribute to understanding better the dynamics of 5mC, 5hmC, 5fC and 5caC. Secondly, as the TETs and 5hmC represent just a piece of the bigger epigenetic puzzle, deciphering the interplay between DNA methylation, DNA hydroxymethylation and histone modifications, is crucial to understanding the processes that ultimately drive or suppress gene expression. And finally, but most importantly, the knowledge we gather along the way should be translated into the development of "smarter", more sophisticated (epigenetic) treatments, than the ones we can currently offer to patients with myeloid malignancies.

REFERENCES

1. Yang H, Liu Y, Bai F, Zhang JY, Ma SH, Liu J, Xu ZD, Zhu HG, Ling ZQ, Ye D, et al. Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. Oncogene 2013; 32:663-9.
2. Wielscher M, Liou W, Pulverer W, Singer CF, Rappaport-Fuerhauser C, Kandioler D, Egger G, Weinhausel A. Cytosine 5-Hydroxymethylation of the LZTS1 Gene Is Reduced in Breast Cancer. Translational oncology 2013; 6:715-21.
3. Jawert F, Hasseus B, Kjeller G, Magnusson B, Sand L, Larsson L. Loss of 5-hydroxymethylcytosine and TET2 in oral squamous cell carcinoma. Anticancer research 2013; 33:4325-8.
4. Orr BA, Haffner MC, Nelson WG, Yegnasubramanian S, Eberhart CG. Decreased 5-hydroxymethylcytosine is associated with neural progenitor phenotype in normal brain and shorter survival in malignant glioma. PloS one 2012; 7:e41036.
5. Gambichler T, Sand M, Skrygan M. Loss of 5-hydroxymethylcytosine and ten-eleven translocation 2 protein expression in malignant melanoma. Melanoma research 2013; 23:218-20.
6. Yang Q, Wu K, Ji M, Jin W, He N, Shi B, Hou P. Decreased 5-hydroxymethylcytosine (5-hmC) is an independent poor prognostic factor in gastric cancer patients. Journal of biomedical nanotechnology 2013; 9:1607-16.
7. Liu WR, Tian MX, Jin L, Yang LX, Ding ZB, Shen YH, Peng YF, Zhou J, Qiu SJ, Dai Z, et al. High expression of 5-hydroxymethylcytosine and isocitrate dehydrogenase 2 is associated with favorable prognosis after curative resection of hepatocellular carcinoma. Journal of experimental & clinical cancer research : CR 2014; 33:32.
8. Ko M, An J, Bandukwala HS, Chavez L, Aijo T, Pastor WA, Segal MF, Li H, Koh KP, Lahdesmaki H, et al. Modulation of TET2 expression and 5-methylcytosine

oxidation by the CXXC domain protein IDAX. Nature 2013; 497:122-6.

9. Kojima T, Shimazui T, Hinotsu S, Joraku A, Oikawa T, Kawai K, Horie R, Suzuki H, Nagashima R, Yoshikawa K, et al. Decreased expression of CXXC4 promotes a malignant phenotype in renal cell carcinoma by activating Wnt signaling. Oncogene 2009; 28:297-305.
10. Zhang P, Huang B, Xu X, Sessa WC. Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. Biochemical and biophysical research communications 2013; 437:368-73.
11. Song SJ, Ito K, Ala U, Kats L, Webster K, Sun SM, Jongen-Lavrencic M, Manova-Todorova K, Teruya-Feldstein J, Avigan DE, et al. The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. Cell stem cell 2013; 13:87-101.
12. Cheng J, Guo S, Chen S, Mastriano SJ, Liu C, D'Alessio AC, Hysolli E, Guo Y, Yao H, Megyola CM, et al. An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis. Cell reports 2013; 5:471-81.
13. Huang H, Jiang X, Li Z, Li Y, Song CX, He C, Sun M, Chen P, Gurbuxani S, Wang J, et al. TET1 plays an essential oncogenic role in MLL-rearranged leukemia. Proceedings of the National Academy of Sciences of the United States of America 2013; 110:11994-9.
14. Gelmetti V, Zhang J, Fanelli M, Minucci S, Pelicci PG, Lazar MA. Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. Molecular and cellular biology 1998; 18:7185-91.
15. Wang J, Hoshino T, Redner RL, Kajigaya S, Liu JM. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proceedings of the National Academy of Sciences of the United States of America 1998; 95:10860-5.
16. Liu S, Shen T, Huynh L, Klisovic MI, Rush LJ, Ford JL, Yu J, Becknell B, Li Y, Liu C, et al. Interplay of RUNX1/MTG8 and DNA methyltransferase 1 in acute myeloid leukemia. Cancer research 2005; 65:1277-84.
17. Hatlen MA, Wang L, Nimer SD. AML1-ETO driven acute leukemia: insights into pathogenesis and potential therapeutic approaches. Frontiers of medicine 2012; 6:248-62.
18. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 2011; 473:343-8.
19. Putiri EL, Tiedemann RL, Thompson JJ, Liu C, Ho T, Choi JH, Robertson KD. Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. Genome biology 2014; 15:R81.
20. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic acids research 2011; 39:5015-24.
21. Tan L, Shi YG. Tet family proteins and 5-hydroxymethylcytosine in development and disease. Development 2012; 139:1895-902.
22. Song CX, Yi C, He C. Mapping recently identified nucleotide variants in the genome and transcriptome. Nature biotechnology 2012; 30:1107-16.
23. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Munzel M, Wagner M, Muller M, Khan F, et al. Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. Cell 2013; 152:1146-59.
24. Iurlaro M, Ficz G, Oxley D, Raiber EA, Bachman M, Booth MJ, Andrews S, Balasubramanian S, Reik W. A screen for hydroxymethylcytosine and formylcytosine binding proteins

suggests functions in transcription and chromatin regulation. Genome biology 2013; 14:R119.

25. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 2012; 149:1368-80.
26. Otani J, Kimura H, Sharif J, Endo TA, Mishima Y, Kawakami T, Koseki H, Shirakawa M, Suetake I, Tajima S. Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. PloS one 2013; 8:e82961.
27. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, Cheng X. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic acids research 2012; 40:4841-9.
28. Cartron PF, Nadaradjane A, Lepape F, Lalier L, Gardie B, Vallette FM. Identification of TET1 Partners That Control Its DNA-Demethylating Function. Genes & cancer 2013; 4:235-41.

Chapter 6

Scope of this Thesis

In 2009 novel mutations in the TET2 gene were reported in myelodysplastic syndromes (MDS) and myeloproliferative neoplasms (MPN). TET2 belongs to a family of proteins, including also TET1 and TET3. The TETs belong to the group of 2-oxoglutarate and Fe-dependent dioxygenases and catalyze the sequential oxidation of 5-methylcytosine 5mC) into 5-hydroxymethyl- (5hmC), 5-formyl- (5fC) and finally 5-carboxyl-cytosine (5caC). 5caC is recognized by the cell's base-excision repair machinery and replaced by an unmethylated cytosine. Essentially, the discovery of the TET proteins marked a new era in the area of active DNA demethylation, a process that had previously remained controversial and elusive. Since epigenetic deregulation is a hallmark of cancer, we wanted to further investigate the significance of TET2 mutations, both clinically and functionally. Approximately 30% of MDS patients eventually progress to acute myeloid leukemia (AML). In **Chapter 2** of this thesis, we determined the clinical and biological significance of TET2 aberrations in a well-characterized cohort of younger AML patients (age ≤60). We found that TET2 mutations affect approximately 8% of the patients, but more importantly, that they are a marker of poor prognosis. TET2 mutations are mutually exclusive of IDH defects, but frequently co-occur with both NPM1 and DNMT3a mutations. In addition, patients who carry both TET2 and NPM1, and TET2 and DNMT3a aberrations have a worse overall survival compared to NPM1 or DNMT3a single mutations. Our next goal was to determine what the biological effect of TET2 mutations was. Overexpression of six mutant TET2 sequences in cell lines showed that all TET2 mutations are loss-of-function, but are not dominant negative. Given that loss-of-function TET2 aberrations had a negative impact on OS of patients we further investigated whether low TET2 expression in TET2 wild type AML patients would affect outcome. We found that low TET2 expression, which is not caused by TET2 defects, may indeed result in inferior survival in AML.

Following up on this study, in **Chapter 3,** we measured 5hmC levels in the same cohort of AML patients as to find out whether 5hmC levels have prognostic value. First we compared levels of 5hmC in healthy vs. AML patient cells. Notably, healthy cells showed a narrow range of 5hmC values, whereas patient cells had a much wider distribution of 5hmC levels, already suggesting that deregulation of DNA hydroxymethylation may be a distinct feature of AML. Within the patient group, we identified three subpopulations based on 5hmC levels, namely low, intermediate and high. TET2 and IDH1/2 mutations both result in reduced levels of 5hmC. Most of the patients with low 5hmC in our study did in fact carry either a TET2, or an IDH aberration. TET2 mutations are known to have a negative impact on survival, but the effect of IDH defects remains unclear. The significance of low 5hmC levels on OS in the context of TET2 and IDH mutations should be investigated in larger patient cohorts. Interestingly, high 5hmC levels were an independent indicator of poor prognosis in AML. In addition, they correlated with MLL and AML1-ETO

aberrations. Altogether, this study suggests that aberrant 5hmC levels can be independent of TET2 and IDH mutations, but still have a significant impact on OS of AML patients.

Although it is clear that 5hmC is an intermediate along the DNA demethylation pathway, many questions regarding its function(s) remain unanswered at present. The high levels of 5hmC in some tissues, such as the brain, as well as the existence of 5hmC-specific readers, have led to the hypothesis that the "sixth" DNA base may in fact be a functional mark, and not simply a demethylation by-product. However, the evidence presented in **Chapter 4** of this thesis does not support this hypothesis. Unlike DNA methylation levels which remained stable during DNA replication, DNA hydroxymethylation levels dropped to approximately 80% in late S and G2/M phases of the cell cycle and were restored during G0/G1. The reduction of 5hmC levels was not related to TET, as neither mRNA, nor protein levels of TET1, TET2 and TET3, changed during various stages of the cell cycle. Since 5hmC can only be formed based on the pre-existence of 5mC in the genome, we conclude that DNA hydroxymethylation is secondary to DNA methylation.

Chapter 7

English Summary

Blood is a highly specialized tissue consisting of different types of cells, namely erythrocytes (red blood cells), leukocytes (white blood cells) and thrombocytes (platelets), suspended in plasma. Red blood cells are the most abundant and their function is to transport oxygen in the body. White blood cells represent the immune system and defend the body against infections and foreign material, whereas platelets are responsible for blood clotting. All mature blood cells originate from pluripotent hematopoietic stem cells - special cells, which are capable of differentiating (maturing) into multiple cell types. Hematopoiesis (from Ancient Greek: *haima* blood and *poiesis* to make), or the formation of all mature blood cells, before they enter the bloodstream, is a complex process, which takes place in the bone marrow.

Sometimes, due to mistakes in the DNA sequence, hematopoiesis is impaired. Immature blood cells become blocked at a certain stage in their maturation and cannot progress further. In addition, they lose their ability to control cell division and start expanding uncontrollably. As a result, the bone marrow and/or blood become overpopulated by non-functioning, immature blood cells - a cancer of the blood (leukemia). Generally, depending on the blood cell type, and stage at which cell differentiation is arrested, different types of leukemia exist. Patients with leukemia suffer from anemia, susceptibility to infections and inability to clot, due to the lack of red and white blood cells, and platelets.

In the past 50 years, especially after the discovery of DNA structure in 1953, researchers have made significant progress in understanding the human genome - the genetic material of an organism. Every cell in our body contains DNA in its nucleus. DNA consists of 4 building blocks: A (adenine), (G) guanine, C (cytosine) and T (thymine). A gene is a short stretch of DNA, which contains the instructions for the synthesis of a protein. Proteins are essential to the proper functioning of cells, tissues and ultimately the whole organism.

Mistakes in a particular DNA sequence, for instance the substitution of an A for a G, are called mutations. Leukemia is caused by various genetic mutations. Our increasing knowledge of how these mutations affect the corresponding proteins and disrupt cellular processes has led to more refined leukemia classifications, more precise diagnosis and prognostic scoring of patients, and will eventually translate into better treatment strategies. Genetic testing already gives physicians the opportunity to offer more tailored treatment, based on the specific set of mutations that a patient carries. Certain mutations give a better prognosis and can be approached with less intensive cures, whereas others have a worse prognosis. Patients with a poorer prognosis and increased risk of relapse are treated more aggressively.

In 2009, our group found a novel mutation in a gene that had never been described before. The gene is called *TET2*, and although at the time we didn't know that, it functions as an epigenetic regulator. Since every single cell in our bodies

carries that same genetic information, there have to be mechanisms which regulate gene expression. Essentially this means that certain genes, which are active in liver cells for instance, should not be expressed in neurons, and vice versa. These highly specific patterns of gene expression are possible due to the existence of the "above genetic" code, i.e. the epi-genetic code. The most famous epigenetic modification is DNA methylation, which entails the addition of a methyl group to cytosine (C => 5mC), one of the building blocks of DNA. Stretches of DNA which contain methylated cytosines signify repression of gene expression. The *TET2* gene converts 5-**methyl**cytosine (5mC) into 5-**hydroxymethyl**cytosine (5hmC), and further to 5-**formyl** and 5 **carboxyl**-cytosine. The formyl and carboxyl derivatives are then excised by specific enzymes and replaced by C (cytosine), thus reverting gene silencing. Regulation of gene expression is an extremely complex process, and the TET2 protein turned out to be very important player in it.

In **Chapter 2** of this thesis, we determined the prognostic and biological significance of *TET2* mutations in a specific type of leukemia, called acute myeloid leukemia (AML). AML is the most common type of acute leukemia in adults, and its incidence increases with age. We wanted to determine how *TET2* mutations affect prognosis of AML patients. Our study included 357 AML patients and mutations were found in 8% of the patients. Moreover, *TET2* mutations were a marker of poor prognosis. We also measured the levels of expression of the *TET2* gene in patients and discovered that lower levels of *TET2* in AML patients who do not carry a mutation in the gene, also signify worse overall survival. It remains unknown what causes the decrease of *TET2* in these cases.

Although this information already makes a difference for diagnosing and treating a patient with a *TET2* mutation, it still does not explain how the mistake in the gene causes the disease. To answer that question, we made use of a cell line model. Cell lines are immortalized cells which can grow indefinitely in laboratory conditions and are often used in research. Making use of molecular cloning technology, the *TET2* gene was artificially introduced in different cell lines. 5hmC, the product of TET2 activity, was used as readout. Wild type (non-mutated) TET2 was compared to 6 different mutant TET2 proteins. Unlike wild type TET2, which was able to produce 5hmC, none of the 6 mutants did. This shows that *TET2* mutations cause loss-of-function in AML cells which carry the mutation, i.e. mutations found in AML inactivate TET2 function and likely cause disturbances in gene expression.

5hmC is a direct product of the activity of TET2, but it certainly depends on more factors. Some proteins are part of large protein complexes, with each component being responsible for a particular task. Others regulate the expression of proteins by marking them for degradation. This means that proteins, other than TET2, could cause a similar phenotype, i.e. a reduction in 5hmC levels, when they mis-function. For instance mutations in another gene, called IDH, lead to the production of a metabolite, 2-hydroxyglutarate, which inhibits the function of TET2. Patients who

have IDH mutations, but an intact TET2, also present with decreased 5hmC levels. TET2 and IDH mutations are therefore mutually exclusive - they affect the same biological pathway, so they almost never occur together in the same patient. The reason for this is that even if cancer cells do acquire both mutations, that does not provide them bigger growth advantage.

In Chapter 3, we investigated the levels and prognostic value of 5hmC levels in 206 AML patients. 5hmC levels in healthy blood cells varied only in a narrow range, however, 5hmC in AML cells was distributed in a much wider range. Some patients had levels which were in the same scope as healthy cells. Others had lower or higher 5hmC levels. Most of the low 5hmC cases could be explained be the presence of a mutation either in TET2, or in IDH. There were, nonetheless, patients with unexplained low 5hmC levels. It would be very interesting to find out what the cause for that is. More importantly, this study showed that high 5hmC levels are a marker of poor prognosis. Research performed on 5hmC levels in other types of cancer has also indicated that 5hmC levels may be an important biomarker in cancer. Even though, we do see that certain mutations are more prevalent in AML patients with high 5hmC levels, it remains unknown what the precise mechanism is.

Whereas Chapters 2 and 3 focused more on the clinical relevance of *TET2* and 5hmC, Chapter 4 explored aspects of their basic biology.

DNA methylation (5mC), the mark that silences gene expression, is fundamental to cells. It ensures that certain regions of DNA, which need to remain repressed, will be. It is also crucial for genomic imprinting and X-chromosome inactivation. Genomic imprinting is an interesting epigenetic phenomenon leading the parent-of-origin-specific gene expression. We always get one copy of DNA from our mother, and one from our father, meaning that we get two sets of genes. Certain genes are "imprinted" - the copy that is expressed always comes from the same parent, whereas the other always remains silent. DNA methylation mediates that process. X-chromosome inactivation is another fascinating event taking place in the cells of females. Females inherit two copies of the X-chromosome, but to function normally they only need the set of genes located on one X-chromosome. To compensate for that, one X chromosome in the cell is always inactive, or silent. The choice which X-chromosome will be inactivated in any given cell in the body is random. DNA methylation makes sure that one X-chromosome in each cell remains silent and that effect lasts over the next cell generations, the effect is heritable.

When cells divide, the need to first double, i.e. replicate, their DNA, so that each daughter cells gets an identical set of genetic information. During DNA replication, the DNA double helix unwraps and each strand is used as a template, so that after completion of cell division one strand comes from the mother cell and one strand, which is identical, is newly synthesized. In theory, DNA replication poses a problem for DNA methylation, as the daughter DNA strand would contain a normal cytosine, instead of a methylated one (C instead of 5mC) at the end of

the process. This would be detrimental to the cell. The examples in the previous paragraph demonstrate the importance of epigenetic memory. Therefore, cells have developed a mechanism to prevent that from happening. A specific enzyme called DNA-methyltransferase 1, DNMT1, makes sure that cytosines on the newly synthesized DNA get methylated already during DNA replication.

In Chaper 4 of this thesis we asked whether 5hmC, like 5mC, is inherited during DNA replication. This study not only aimed at answering a basic scientific question about 5hmC, but also at shedding light on the discussion whether DNA hydroxymethylation is a stable epigenetic mark, or simply an intermediate in the pathway that reverses gene silencing caused by DNA methylation. We measured both 5mC (as a control) and 5hmC. Due to the mechanism for 5mC transmission during DNA replication, its levels remained unchanged at different phases of the cell cycle. In contrast, the levels of 5hmC decreased during DNA replication, showing that this epigenetic mark is not sustained in a similar manner, as 5mC. This leads us to believe that, although 5hmC is an important component of the DNA (de)methylation pathway, it is not a stand-alone epigenetic mark, but rather an intermediate in the process. It would be extremely interesting to investigate in depth what factors dictate the formation of 5hmC at given DNA sites and how that relates to the life cycle of the cell.

It is my hope that future research will unravel the many yet unanswered questions about TET2 and 5hmC, but most of all, that this will result in the development of targeted epigenetic therapies for patients with defects in the DNA (de)methylation pathway.

Chapter 8

Nederlandse Samenvatting

Bloed is een zeer gespecialiseerd weefsel dat bestaat uit verschillende soorten cellen, namelijk erythrocyten (rode bloedcellen), leukocyten (witte bloedcellen) en thrombocyten (bloedplaatjes). Deze zijn opgelost in bloedplasma. Rode bloedcellen zijn het meest aanwezig en hun functie is het vervoeren van zuurstof in het lichaam. Witte bloedcellen vertegenwoordigen het immuunsysteem en verdedigen het lichaam tegen infecties en vreemd materiaal. Bloedplaatjes zijn betrokken bij bloedstolling. Alle volwassen bloedcellen zijn afkomstig uit pluripotente hematopoëtische stamcellen. Dit zijn speciale cellen die in staat zijn te differentieren (groeien) naar verschillende soorten cellen. Hematopoëse (uit het oud-grieks: *haima* is bloed en *poiesis* is maken), of het vormen van alle soorten bloedcellen, voordat ze het bloedsomloop toetreden, is een complex proces dat plaats vindt in het beenmerg.

Er zijn voorvallen waarbij hematopoëse verzwakt door fouten in de DNA sequentie. Onvolgroeide bloedcellen worden geblokkeerd in een van hun groeifases en zijn niet in staat verder te ontwikkelen. Verder zijn ze niet meer in staat hun celdeling te controleren en breiden zich ongeremd uit. Als resultaat wordt het beenmerg en/of bloed overvloed door niet-functionerende onvolgroeide bloed cellen – er is sprake van bloedkanker (leukemie). In het algemeen, afhankelijk van het type bloedcel en de groeifase waar het differentiëren gestopt is, bestaan er verschillende soorten leukemie. Patiënten met leukemie lijden aan anemie, kwetsbaarheid voor infecties en het onvermogen tot bloedstolling. Dit komt door het tekort aan rode bloedcellen, witte bloedcellen en bloedplaatjes.

In de afgelopen 50 jaar, vooral na de ontdekking van de DNA structuur in 1953, hebben onderzoekers aanzienlijke vooruitgang geboekt in het begrijpen van het humaan genoom – het genetisch materiaal van het menselijk organisme. Elke cel in ons lichaam bevat DNA in de celkern. DNA bestaat uit 4 bouwstenen: A (adenine), G (guanine), C (cytosine) en T (thymine). Een gen is een kort stukje DNA dat de instructies bevat voor de synthese van een eiwit. Eiwitten zijn essentieel voor het functioneren van cellen, weefsels en uiteindelijk het hele organisme.

Fouten in een DNA sequentie, bijvoorbeeld het vervangen van een A door een G, worden mutaties genoemd. Leukemie wordt veroorzaakt door verschillende genetische mutaties. Onze groeiende kennis van hoe deze mutaties hun bijbehorende eiwitten beïnvloeden en cellulaire processen verstoren heeft geleid tot verfijnde leukemie classificaties en preciezere diagnoses van patiënten. Dit zal uiteindelijk ook leiden tot betere strategiën voor behandeling. Genetische tests helpen artsen behandelingen op maat te bereiden, gebaseerd op de specifieke set van mutaties die een patiënt draagt. Bepaalde mutaties geven een betere prognose en kunnen benaderd worden met minder intensieve behandelingen. Patiënten met een slechtere prognose en verhoogde kans op terugval worden agressievere therapiën behandeld.

In 2009 vond onze groep een niewe mutatie in een gen dat nog niet eerder beschreven was. Dit gen is genaamd TET2 en het functioneert als een epigenetische regulator. Aangezien elke cel in onze lichamen dezelfde genetische informatie draagt moeten er mechanismen zijn die genexpressie reguleren. Dit betekent bijvoorbeeld, dat genen die geactiveerd worden in levercellen niet actief moeten worden in neuronen, en vice versa. Deze zeer gespecialiseerde patronen van genexpressie zijn mogelijk door het bestaan van de hogere genetische code, ofwel de epigenetische code. De meest bekende epigenetische modificatie is DNA methylatie. Deze voegt een methyl groep aan cytosine (C -> 5mC), een van de bouwstenen van DNA. Stukken DNA die gemethyleerde cytosines bevatten duiden op onderdrukking van genexpressie. Het TET2 eiwit converteert 5-**methyl**cytosine (5mC) naar 5-**hydromethyl**cytosine (5hmC), en verder naar 5-**formyl**- en 5-**carboxyl**-cytosine. De formyl en carboxyl toevoegingen worden dan met behulp van enzymen weggesneden en weer door cytosine vervangen. Dit brengt het gen dus weer tot expressie/activatie. Regulatie van genepressie is een complex proces en het TET2 eiwit is daar zeer bij betrokken.

In **Hoofdstuk 2** van deze scriptie hebben we de prognostische en biologische significantie van TET2 bepaald. Dit is gedaan voor een specifieke vorm van leukemie, genaamd acute myeloïde leukemie (AML). AML is de meest voorkomende soort leukemie in volwassenen en de incidentie ervan neemt toe met de leeftijd. Wij wilden bepalen hoe TET2 mutaties de prognose van AML patiënten beïnvloeden. Onze studie besloeg 357 AML patiënten en mutaties werden gevonden in 8% van deze groep. Het vinden van een TET2 mutatie duidt aan op een slechte prognose. We hebben de hoeveelheden van het TET2 gen gemeten in de patiënten en vonden dat lagere hoeveelheden van TET2 zonder een mutatie van het gen ook duiden op een slechtere overlevingskans. Het is onbekend wat de verminderde hoeveelheid van TET2 in deze voorvallen veroorzaakt.

Deze informatie is nuttig voor de diagnose en behandeling van een patiënt met de TET2 mutatie. Het legt echter niet uit hoe de fout in het gen de aandoening veroorzaakt. Om deze vraag te beantwoorden maken we gebruik van een cellijn model. Cellijnen zijn onsterfelijke cellen die oneindig lang kunnen doorgroeien in een laboratorioum. Ze worden vaak gebruikt in onderzoek. Met moleculaire kloon technologie was het TET2 gen kunstmatig toegevoegd aan verschillende cellijnen. 5hmC, het product van TET2 activiteit werd uitgelezen. Wild-type (niet-gemuteerde) TET2 werd vergeleken met 6 TET2 mutanten. Het wild-type TET2 was in staat 5hmC te produceren, alle 6 mutanten konden dat niet. Dit toont aan dat AML cellen met TET2 mutaties deactiveren de TET2 functionaliteit en veroorzaken warschijnlijk verstoringen in genexpressie.

5hmC is een directe product van TET2 activiteit, maar het is zeker afhankelijk van meerdere factoren. Sommige eiwitten zijn onderdeel van grote eiwitstructuren, met elke onderdeel verantwoordelijk voor een bepaalde taak. Anderen reguleren

de expressie van eiwitten door ze te markeren voor degradatie. Dit betekent dat naast TET2 ook andere misfunctionerende eiwitten een vergeiljkbaar fenotype kunnen veroorzaken, namelijk een verminderde hoeveelheid van 5hmC. Ter voorbeeld: mutaties in een ander gen, genaamd IDH, leiden to de productie van een metaboliet, 2-hydroxyglutaraat, deze remt het functioneren van TET2. Patiënten met IDH mutaties, maar een intact TET2 tonen ook verminderde hoeveelheden 5mhC. TET2 en IDH muaties zijn dus wederzijds exclusief. Ze beïnvloeden dezelfde biologische pathway, dus ze komen bijna nooit allebei voor in een patiënt. De reden hiervoor is dat kankercellen geen groeivoordeel krijgen zelfs als ze alle twee mutaties bevatten.

In **Hoofdstuk 3** hebben we de hoeveelheden en prognostische waardes van 5hmC onderzocht in 206 AML patiënten. De hoeveelheid 5hmC varieerde minimaal in gezonde bloedcellen. In AML cellen was de variate aan hoeveelheid 5hmC echter veel breder. Sommige patiënten hadden een vergelijkbaar 5hmC niveau als dat van gezonde cellen. Anderen hadden hogere of lagere 5hmC niveaus. De meeste van de lage 5hmC gevallen konden verklaard worden met de aanwezigheid van een mutatie in TET2 of IDH. Er waren echter toch patiënten met een onverklaarbaar laag 5hmC gehalte. Het zou erg interessant zijn om de oorzaak hiervan te achterhalen. Belangrijker is echter dat het onderzoek aanwees dat een hoger 5hmC gehalte een marker is van een slechte prognose. Onderzoek naar 5hmC gehaltes in andere soorten kanker heeft ook aangetoond dat 5hmC een belangrijke biomarker kan zijn in kanker. Ook al zien we dat bepaalde mutaties vaker voorkomen in AML patiënten met hogere 5hmC hoeveelheden, het blijft onbekend wat het precieze mechanisme daaracher is.

Waar Hoofdstuk 2 en 3 meer gefocused waren op de klinische relevantie van TET2 en 5hmC, doet **Hoofdstuk 4** verder kijken op aspecten van hun basale biologie.

DNA methylatie (5mC), het teken dat genexpressie stillegt, is fundamenteel voor cellen. Het zorgt ervoor dat bepaalde gebieden van DNA die onderdrukt moeten worden dat ook daadwerkelijk blijven. Het is ook cruciaal voor genomische inprenting en X-chromosoom inactivatie. Genomische inprenting is een interessant epigenetisch fenomeen. Het houdt in dat genexpressie specifiek behouden blijft volgens het oorspronkelijk gen van de ouder. Wij krijgen altijd een DNA kopie van onze moeder en een van onze vader. We hebben dus twee sets aan genen. Bepaalde genen zijn "imprinted" – de kopie die expressie vertoont komt altijd van dezelfde ouder vandaan, de andere blijft altijd onderdrukt. DNA methylatie fungeert als mediator in dat proces. X-chromosoom inactivatie is nog een fascinerend voorval dat plaats vindt in cellen van vrouwen. Vrouwen erven twee kopiën van van het X-chromosoom. Om daarvoor te compenseren is een X-chromosoom in elke cel altijd inactief en dat effect wordt behouden over toekomstige celgeneraties, het effect is erfelijk.

Wanneer cellen zich delen moeten ze eerst hun DNA verdubbelen, ofwel repliceren, zodat elke dochtercel een identieke set aan genetische informatie ontvangt. Tijdens DNA replicatie wordt de DNA dubbele helix ontwonden en elke streng wordt gebruikt als een template. Na de celdeling is dan een streng afkomstig van de moedercel en een streng is identiek nieuw gesynthetiseerd. In theorie is DNA replicatie een probleem voor DNA methylatie. De dochter DNA streng zou een normale cytosine bevatten in plaats van een gemethyleerde (C in plaats van 5mC) aan het einde van het proces. Dit zou schadelijk zijn voor de cel. De voorbeelden in de vorige paragraaf demonstreren het belang van epigenetische geheugen. Cellen hebben dus een mechanisme ontwikkeld om dit probleem te voorkomen. Een specifiek enzym genaamd DNA-methyltransferase 1, DNMT1, zorgt ervoor dat cytosines op nieuw gesynthetiseerd DNA gemethyleerd worden. Dit gebeurt al tijdens DNA replicatie.

In **Hoofdstuk 4** van deze scriptie hebben we de vraag gesteld of 5hmC, net als 5mC, geërfd wordt tijdens DNA replicatie. Dit onderzoek is niet alleen gericht om een basale wetenschappelijke vraag over 5hmC the beantwoorden, maar ook om zicht te werpen op de discussie of DNA hydroxymethilatie een stabiele epigenetische mark is, of alleen een intermediair in de pathway dat het onderdrukken van genen omzet door DNA methylatie. We hebben zowel 5mC (als controle) als 5hmC gemeten. Door het mechanisme voor 5mC transmissie tijdens DNA replicatie zijn de gehaltes onveranderd gebleven tijdens de verschillende fases van de celcyclus. De gemeten niveaus van 5hmC waren in contrast verminderd tijdens DNA replicatie, wat aantoonde dat deze epigenetische mark niet op een vergelijkbare manier behouden wordt, zoals 5mC. Dit laat ons geloven dat, ook al is 5hmC een belangrijk component van de DNA (de)methylatie pathway, het is geen epigenetiche mark op zich, maar eerder een tussenstap in het proces. Het zou erg interessant zijn om nader te onderzoeken welke factoren de formatie van 5hmC op specifieke DNA locaties dicteren en hoe dit gerelateerd is aan de levenscyclus van de cel.

Het is mijn hoop dat toekomstig onderzoek vele van deze vragen over TET2 en 5hmC zal beantwoorden, maar vooral dat het resultaat zal bijdragen aan ontwikeling van gerichte epigenetische therapiën voor patiënten met defecten in de DNA (de)methylatie pathway.

LIST OF PUBLICATIONS

Acquired mutations in TET2 are common in myelodysplastic syndromes
Langemeijer SM*, Kuiper RP*, Berends M, Knops R, **Aslanyan MG**, Massop M, Stevens-Linders E, van Hoogen P, van Kessel AG, Raymakers RA, Kamping EJ, Verhoef GE, Verburgh E, Hagemeijer A, Vandenberghe P, de Witte T, van der Reijden BA, Jansen JH.
Nat Genet. 2009 Jul;41(7):838-42. doi: 10.1038/ng.391. Epub 2009 May 31

TET proteins in malignant hematopoiesis
Langemeijer SM, **Aslanyan MG**, Jansen JH.
Cell Cycle. 2009 Dec 15;8(24):4044-8. Epub 2009 Dec 5. No abstract available.
PMID: 19923888 [PubMed - indexed for MEDLINE

TET2 mutaties komen frequent voor bij myeloïde maligniteiten
S.M.C. Langemeijer, **M.G. Aslanyan** en J.H. Jansen
Ned Tijdschr Hematol 2010;7:131, 2010

The mitotic spindle protein SPAG5/Astrin connects to the Usher protein network postmitotically
Kersten FF, van Wijk E, Hetterschijt L, Bauβ K, Peters TA, **Aslanyan MG**, van der Zwaag B, Wolfrum U, Keunen JE, Roepman R, Kremer H.
Cilia. 2012 Apr 25;1(1):2. doi: 10.1186/2046-2530-1-2.

Characterization of acute myeloid leukemia based on levels of global hydroxymethylation
Kroeze LI*, **Aslanyan MG***, van Rooij A, Koorenhof-Scheele TN, Massop M, Carell T, Boezeman JB, Marie JP, Halkes CJ, de Witte T, Huls G, Suciu S, Wevers RA, van der Reijden BA, Jansen JH.
Blood. 2014 Jul 1. pii: blood-2013-08-518514

Clinical and biological impact of TET2 mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial
Aslanyan MG, Kroeze LI, Langemeijer SM, Koorenhof-Scheele TN, Massop M, van Hoogen P, Stevens-Linders E, van de Locht LT, Tönnissen E, van der Heijden A, da Silva-Coelho P, Cilloni D, Saglio G, Marie JP, Tang R, Labar B, Amadori S, Muus P, Willemze R, Marijt EW, de Witte T, van der Reijden BA, Suciu S, Jansen JH.
Ann Hematol. 2014 Aug;93(8):1401-12. doi: 10.1007/s00277-014-2055-7. Epub 2014 Mar 29

* shared co-first authorship

CURRICULUM VITAE

Mariam Aslanyan was born on January 6th, 1983 in Sofia, Bulgaria. She attended a foreign languages-oriented high-school, where she focused on English language and literature. After graduating with a summa cum laude, Mariam continued her education at the Faculty of Biology of the Sofia University "St. Kliment Ohridski" in 2001. As she had developed an aptitude for the life sciences, she chose to major in Biology and Chemistry. In 2005, she obtained her BSc degree (cum laude) and subsequently enrolled in a master's program at the Faculty of Chemistry, entitled "Spectral and Chromatographic Methods for Analysis". During that time she became deeply interested in Infrared Spectroscopy. After a year at the Faculty of Chemistry, Mariam realized that her heart lies with molecular biology, rather than the chemical sciences. In 2006, she applied and was admitted to the highly competitive Molecular Mechanisms of Disease (MMD) master's program at the Nijmegen Centre for Molecular Life Sciences of the Radboud University in Nijmegen, the Netherlands. She received scholarships both from the Radboud University, as well as the Huygens Program. During her MMD studies, Mariam performed two research internships at the UMC St. Raboud in Nijmegen. From January till November 2007, she studied genetic mutations in patients with myelodysplastic syndromes (MDS), under the supervision of Dr. Saskia Langemeijer and Prof. Dr. Joop Jansen, at the Hematology Lab. Afterwards, from December 2007 till July 2008, Mariam performed research on the topic of Usher Syndrome in the group of Prof. Dr. Hannie Kremer at the Department of Human Genetics, under the supervision of Dr. Ferry Kersten and Dr. Erwin van Wijk. Upon completion of this internship, she obtained her MSc diploma (cum laude) in August 2008. In September that same year, she returned to the Hematology Lab to begin her PhD studies following up on an exciting discovery that was made during her first research internship – namely, the sequencing of the first TET2 mutation in an MDS patient. The focus of her PhD project was the TET2 gene/protein.

Chapter 9

Dankwoord/Acknowledgements

Милички мои, мами, тате и Кевчо, започвам с вас! Закъде бих била без вас... Вие сте всичко за мен!!! Независимо от разстоянието и времето ние винаги сме заедно (нека да благодарим на всички иновативни технологии за тази привилегия), което ми помогна да оцелея далеч от вас през всички тези години. Това приключение и поредният му успех дължа на вас и на никой друг! Няма думи, с които да опиша колко съм благодарна, че ви има. Няколко реда са крайно недостатъчни, за да изкажат какво означавате за мен! Дълбоко оценявам вашата неспирна подкрепа, тя е извор на сила за мен. Много ви о.!!!

Чичо, лельо, Они, малко хора на този свят са толкова привилегировани, колкото съм аз! Моето семейство е голямо и прекрасно! Няма по-голямо удоволствие от това да се прибера вкъщи и празникът да се мести от 7-мия към 6-ия етаж и обратно. Много ви обичам!!!

Did you just skim to get to this part?! Gotcha! That's ok though, I know this is THE most anticipated chapter in one's thesis!

At the end of this long, winding road, with all of its ups and downs, there are a lot of people I would like to acknowledge for their direct, or indirect, contribution to this thesis.

Dear Joop, it has been almost nine years since this (slightly scared) MMD student first set foot at the Hematology Lab in late November of 2006. What a long way we have come since then! Admittedly, it was Bert who really inspired me to request an internship at the MHU, but due to circumstances, I ended up starting my internship under yours and Saskia's supervision. PCRs paid off and we made a phenomenal (at least for Master's student standards!) discovery. A gene noone had ever reported before turned out to be frequently mutated in patients with MDS. You must have seen potential in me back then, because following this, you fully supported my wish to write a NWO Toptalent grant proposal and even though, eventually, I did not get it, you hired me as a PhD student on that very same project nonetheless. Little did we know at the time that TET2 would become such a celebrity within the research community, causing quite some let downs on my side. Thank you for always noticing the moments when I really needed a pep talk! One of the best qualities you possess as a supervisor is your humane touch and care that extends beyond the work place. And, I must say, despite the hardships, new ideas always kept flowing in our work discussions...at time, perhaps, in too many different directions. Thank you for your "helicopter" view! Whenever I seemed to be drowning in the details, you were always there to pull me back and show me the bigger picture! I hope to someday be able to sift out the important out of the ocean of irrelevant, with the ease that you seem to do it with! I admire and respect so many things about you as a supervisor, but my fondest memories of you will always be the moments when I used to knock on your door late in the afternoon with the question, "Do you want to see something

cool?!", and then you would turn to me, your eyes would light up, and you'd reply, "ALWAYS!".

Dear Bert, you are the most inspired researcher I have ever met!!! I still remember, clear as day, my first encounter with you. I was an MMD student and you were leading one of our master classes. The topic was CML, and we discussed some articles you had given us the task to read over lunch about ABC transporters on CD34+ cells that pump out Imatinib, causing resistance to the drug. You were telling all of this with such zest, as if it were the most interesting thing on the planet! Then and there, I knew, I have to do my first internship in your group! You couldn't offer me an internship position though, and after my initial disappointment, it was arranged that I start with Joop and Saskia instead. Later, I found out that your passion far outreaches ABC transporters and, basically, covers most things science (is there a science you are not passionate about?). Thank you for the wonderful discussions, ideas, advice, but most of all your contagious enthusiasm! I hope that you will always be able to do research!

Lieve Thess, where do I even begin?! We only met at the very start of my PhD, but like me, you were an "old" face at the lab...albeit coming from the enemy lines haha! Whatever passion you had for Triad though (sorry Triad folks!), was quickly re-directed to the TET2 front. You were the Robin to my Batman, the salt to my pepper, the best co-pilot I could have wished for! Through the years we fought, side-by-side, to unravel the mysteries of our baby - TET2. And our baby sure wasn't an easy one! But we prevailed, and in spite of this love-hate relationship we had with our child, we did manage to make a thesis out of it! This is not only my achievement, Thess, but it is yours too! I am so glad that besides being colleagues, we became good friends too. Thank you for introducing Nederlandse Vrijdag, for the lively scientific discussions, for our long conversations about life and the future!

Lieve Sas (L.), do you remember the afternoon when I came into your office to show you the nonsense mutation I had just found in TET2?! It was the first time after months and months of fruitless PCRs that I finally had something, and I was not quite sure whether it wasn't merely my imagination playing tricks on me, so I had to come and show you. We were both exhilarated!!!... It seems to me like it was only yesterday, but in fact, it has been nine years since...Thanks to you, MDS patients were not just an abstract notion in my head, a source of DNA material and cells. You taught me to look past the cold science and appreciate the fact that there are actual people involved, and that we are doing this research for them. At the start of my PhD, I found it very difficult to evolve out of our supervisor -student relationship, and, in a way, perhaps I never quite did. Your professionalism and dedication have always evoked my admiration and respect for you. Thank you for allowing me not only in your research project, but also in your life! I cannot wait to meet him/her?!?!

Lieve Leeuwtje, dear Leonie, you joined the TET2 forces later, but quickly became an indispensable wheel in the carriage. You joined the group in a way

that was so natural and effortless, and I have no doubts you owe that to your easy-going and lively personality. I enjoyed working with you, just as much as I enjoyed going to the SYTYCD show together. Thank you for sharing the TET2 laughs, the TET2 tears and all the in-betweens! I very much look forward to reading your thesis! Only, I fear leaving you alone with 'The Boys" in the AIO kamer may have had some lasting consequences on your English vocabulary!

Dear Ruth, thank you for contributing you extensive experimental skills and knowledge to the TET2 project! You are now busy with some 5-hmC-related work and I am very curious what you will discover!

Lieve MHD mensen: Marion, Patric, Ellen, Ad, Evelyn, Louis en Laura, veel van jullie waren op een of ander moment direct betrokken bij het TET2 project. Wat hebben we veel samen gesequenced!!! Jullie zijn net een kleine familie en ik volede me altijd op mijn gemak bij de diagnostiek. Bedankt voor jullie inzet en enthusiasme! Het was mij een plezier om jullie te leren kennen!

My dear Minh, you were the first (MMD) student I supervised and you turned out to be so much more than a talented student. I really enjoyed doing TET2 WB and transductions together with you, whilst throwing Friends quotes at each other every few minutes! We have remained good friends and I cherish that! Could I be any clearer hehe!?!?! I have no doubt you will bring your own PhD project to an excellent end!

Dear Sudeep, your task was not an easy one! Cell cycle synchronizations, endogenous TET2 WB, testing different lysing buffers, were among the few experiments you had on your plate during your very first MMD internship. It can be utterly overwhelming, but you did it with pleasure and never complained, even when at first there was not much to be seen on your blots. Thank you for the hard work and enthusiasm! Good luck with your PhD!

My AIO room homies, Davide, Pedro, Jimmy and Nicco, how I have enjoyed your company in the past years!!! Davide, bello, noone can make me laugh by just opening the door to a room, the way you can! You are as much a talented researcher, as you are a wonderful friend! There is so much I would like to say to you, so much I'd like to thank you for, but scribbling a few lines in an acknowledgments section doesn't feel enough. Thank you for always being there for me (also on behalf of Naso Gigio!)! Jim, J-man, it has been an honour and a privilege for me to be your ass buddy! We have been through so many ups and downs in the past few years, but we shall prevail; even in our non-scientific endeavors we remain a great team. I hope that you will not let anything stop you in achieving your goals in life! Thank you for being you! Davide and Jimmy, I am so honored that you have agreed to be my paranimfen! Pedro, my dearest Pjotr, never have I imagined that one of my best friends in this life will be Jesus/the Wolverine himself! Thank you for our Sunday evening tradition, thank you for all of our splendid scientific and non-scientific discussions, our visits to the Marienburg bibliotheek, and, for playing the smallest

violin in the world (when the occasion called for it)! Nicco, although you joined the lab almost at the very end of my lab adventures, I still feel like I've known you for years! Noone tells a made-up story and makes it sound so perfectly believable like you do...or, perhaps I am too gullible?! Thank you for the great times!

Lieve Goric, sweetie, only one month spent under your supervision in November 2007 learning ChIP, convinced me that you are all kinds of amazing! You are not only a true lady, but also a talented researcher. I am so proud of you for choosing to do a PhD! Look how far you've come! Thank you for the appeltaart met moccachino at the Blonde Pater, for our long talks (even when we disagree!), but most of all, for being such a great friend!

Dear Lau, I remember that evening when you and Sas taught me my first sentence in Dutch (De kleren zijn op zolder. *wink, wink*), and the day we were at the Electrophoresis room at the old lab, when you and Sas finally admitted to being a couple (which I suspected all along). So many great memories! You kind of felt sorry for my ever-failing PCRs and decided to take me under your wing to learn some more techniques. Sadly you left the lab soon after I started my PhD, but as far as I am concerned, you were utterly missed! Thank you for always being so kind to me!

Dear Sas (B), Sylvie, Maaike and Lei Lei, I may not have been very close to you over the years, but I will remember each and every one of you. Sas, you were always there to answer my questions and help me out with my experiments, reagents, or just simple orders. Sylvie, how are you so good in everything you do?! Maaike, thank you for your encouraging words and sweet cards! Chocolate can solve any problem, right! Lei Lei, it always amazed me how much courage and kindness there is in such a tiny girl, such as yourself. Thank you for the good times at the Hematology Lab!

Dear Arno and Ron, thank you for working with us on the 5(h)mC measurements! Your contribution to this thesis has been indispensable and I deeply appreciate it!

Dear Rob, thank you for helping me/us with the cell cycle sorting experiments! You were always ready to push your boundaries (and those of the cell sorter!) in order to get the appropriate populations of cells. Thank you for the enthusiasm!

Lieve, lieve Nelleke, bedankt voor al je hulp tijdens mijn promotie!!! Je bent net als een superheld die altijd in de schaduw blijft, maar zonder jou zijn wij allemaal verloren! Jij bent een schat!

My dearest family, Nin&Mo, Shankar&Helene, Manu, Ruggero, Vinc&Dani, Jordan, there are not enough words in my vocabulary to express what you mean to me. I owe you some of the best times of my life! It breaks my heart that so many of you moved away over the past few years, but then again it means I have family in Sweden, Italy, Canada and Barcelona! Not too shabby!!!

Lieve Ali, Zoran, Damir en Marion, jullie hebben mij verwelkomd in de familie Popovic met open armen (en harten). Ik voelde me na een korte tijd een van jullie!

Wat een geluksvogel ben ik, om een familie te hebben gevonden in dit vreemde land, net zoals jullie! Hvala!

And, before I conclude this long (long) acknowledgments section, I turn to you Vanja. Baby (a word only you know all the nuances of!), you mean everything to me! You are my past, my present and my future... and I will not even attempt to thank you in the few short lines that I have left here. You have made my life brighter, happier and more meaningful, and for that I should thank you every day!